

Современный
Гуманитарный
Университет

Дистанционное образование

1001.02.01; 1

Рабочий учебник

Фамилия, имя, отчество _____

Факультет _____

Номер контракта _____

ВЫЧИСЛИТЕЛЬНАЯ МАТЕМАТИКА

ЮНИТА 2

ЧИСЛЕННЫЕ МЕТОДЫ МАТЕМАТИЧЕСКОГО АНАЛИЗА
И РЕШЕНИЯ СИСТЕМ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ

Москва 2002

Разработано В.Н. Варапаевым, д-ром физ.-мат. наук, профессором

Рекомендовано Министерством
общего и профессионального
образования Российской Федерации
в качестве учебного пособия для
студентов высших учебных заведений

КУРС: ВЫЧИСЛИТЕЛЬНАЯ МАТЕМАТИКА

Юнита 1. Численное решение систем линейных и нелинейных уравнений.

Юнита 2. Численные методы математического анализа и решения систем дифференциальных уравнений.

ЮНИТА 2

Изложены основные принципы аппроксимации и интерполяции функций, численного интегрирования и дифференцирования, основные понятия и методы решения разностных и интегральных уравнений, основные понятия и методы численного решения систем дифференциальных уравнений, к которым, как правило, сводятся математические модели.

Для студентов Современного Гуманитарного Университета

Юнита соответствует профессиональной образовательной программе № 1

ОГЛАВЛЕНИЕ

ДИДАКТИЧЕСКИЙ ПЛАН	5
ЛИТЕРАТУРА	6
ПЕРЕЧЕНЬ УМЕНИЙ	7
ТЕМАТИЧЕСКИЙ ОБЗОР	9
1. Аппроксимация и интерполяция функций	9
1.1. Постановка задачи и основные определения	9
1.2. Интерполяция с помощью многочленов	14
1.2.1. Интерполяционный многочлен Лагранжа	15
1.2.2. Интерполяционный многочлен Ньютона	18
1.2.3. Точность и сходимость многочленной интерполяции	21
1.2.4. Использование локальных интерполяций	23
1.3. Равномерные многочленные приближения	25
1.3.1. Использование степенных разложений для вычисления функций	25
1.3.2. Многочлены Чебышева	27
1.3.3. Экономизация степенных рядов с помощью многочленов Чебышева	30
1.4. Среднеквадратичные приближения	33
2. Численное интегрирование	37
2.1. Постановка задачи	37
2.2. Метод прямоугольников	41
2.3. Метод трапеций	44
2.4. Метод Симпсона	48
2.5. Метод Гаусса	51
2.6. Сравнительная оценка методов и способы уточнения решения	58
3. Численное дифференцирование	62
3.1. Постановка задачи	62
3.2. Использование ряда Тейлора	65
3.3. Использование интерполяционных многочленов	69
3.4. Погрешности численного дифференцирования и способы их уменьшения	73
4. Разностные уравнения	75
4.1. Общие понятия	75
4.2. Решение разностных уравнений первого порядка	82
4.3. Решение линейных разностных уравнений второго порядка с постоянными коэффициентами	87
5. Обыкновенные дифференциальные уравнения	89
5.1. Общие понятия	89
5.2. Метод конечных разностей	91
5.3. Решение задачи Коши	95

5.3.1. Основные понятия	95
5.3.2. Метод Эйлера	97
5.3.3. Метод Эйлера с пересчетом	99
5.3.4. Метод Рунге-Кутты	101
5.3.5. Многошаговые методы	105
5.3.6. Сравнительная оценка методов и метод Рунге повышения точности результатов	106
5.3.7. Задача Коши для обыкновенного дифференци- ального уравнения n -го порядка	110
5.4. Решение краевой задачи. Конечно-разностный метод	112
6. Методы решения уравнений в частных производных	120
6.1. Основные понятия и постановка задачи	120
6.2. Описание метода сеток. Аппроксимация, устойчивость, сходимость	125
6.3. Решение параболических задач	130
6.4. Решение эллиптических задач	136
7. Решение интегральных уравнений	140
7.1. Основные понятия	141
7.2. Обзор численных методов	143
ЗАДАНИЯ ДЛЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ	148
ТРЕНИНГ УМЕНИЙ	
ГЛОССАРИЙ*	

* Глоссарий расположен в середине учебного пособия и предназначен для самостоятельного заучивания новых понятий.

ДИДАКТИЧЕСКИЙ ПЛАН

Аппроксимация и интерполяция функций. Постановка задачи. Многочленная интерполяция. Интерполяционные многочлены Лагранжа и Ньютона. Локальная и глобальная интерполяция. Точность и сходимость многочленной интерполяции. Равномерные многочленные приближения. Степенные разложения. Многочлены Чебышева. Экономизация с помощью многочленов Чебышева. Среднеквадратичные приближения.

Численное интегрирование. Методы прямоугольников, трапеций, Симпсона, Гаусса. Сравнительная оценка методов. Способы уточнения решения.

Численное дифференцирование. Использование ряда Тейлора и интерполяционных многочленов. Погрешности численного дифференцирования и способы их уменьшения.

Разностные уравнения. Общие понятия. Решение разностных уравнений первого порядка и линейных разностных уравнений второго порядка с постоянными коэффициентами.

Обыкновенные дифференциальные уравнения. Метод конечных разностей. Решение задачи Коши. Метод Эйлера и его модификации. Метод Рунге-Кутты. Многошаговые методы. Сравнительная оценка методов. Повышение точности результатов методом Рунге. Уравнения n -го порядка, задача Коши. Решение краевой задачи конечно-разностным методом.

Решение уравнений в частных производных. Постановка задачи. Метод сеток, аппроксимация, устойчивость, сходимость. Решение параболических задач. Решение эллиптических задач.

Решение интегральных уравнений. Основные понятия. Обзор численных методов.

ЛИТЕРАТУРА

Базовая

- *1. Турчак Л. И. Основы численных методов: Учебное пособие. М.: Наука, 1987.
- *2. Боглаев Ю. П. Вычислительная математика и программирование: Учебное пособие для студентов вузов. М.: Высшая школа, 1990.
- *3. Бахвалов Н. С., Жидков Н. П., Кобельков Г. М. Численные методы: Учебное пособие. М.: Наука, 1987.

Дополнительная

- 4. Калиткин Н. Н. Численные методы. М.: Наука, 1978.
- 5. Самарский А. А., Николаев Е. С. Методы решения сеточных уравнений. М.: Наука, 1978.
- 6. Форсайт Дж., Малькольм М., Моулер К. Машинные методы математических вычислений. М.: Мир, 1980.
- *7. Ланцош К. Практические методы прикладного анализа: Справочное руководство. М.: Физматгиз, 1961.
- *8. Годунов С. К., Рябенький В. С. Разностные схемы. Введение в теорию. М.: Наука, 1973.
- *9. Хемминг Р. В. Численные методы. М.: Наука, 1968.
- *10. Осиленкер Б. П. Основы математического анализа. Часть II. Юнита 1. М.: СГУ, 1999.
- 11. Варапаев В. Н. Численные методы. Юнита 1. М.: СГУ, 1999.

ПЕРЕЧЕНЬ УМЕНИЙ

№ п/п	Умение	Алгоритмы
1	Использование интерполяционных многочленов для нахождения значений таблично заданных функций	1. По заданному значению x_* определить узловые точки x_i, x_{i+1}, для которых $x_* \in [x_i, x_{i+1}]$. 2. Построить линейный интерполяционный многочлен [см. формулу (1.22)] $L_1(x)$ и вычислить $L_1(x_*)$. 3. Построить квадратный интерполяционный многочлен $L_2(x)$ [см. формулу (1.23)] и вычислить $L_2(x_*)$.
2	Вычисление определенного интеграла численным методом	1. Разделить отрезок интегрирования на n и $2n$ равных частей (с шагами h и $h/2$). 2. По формуле Симпсона вычислить значения интеграла S_n с шагом h и $S_{n/2}$ с шагом $h/2$ [см. формулу (2.25)]. 3. Вычислить уточненное значение интеграла с помощью метода Рунге [см. формулу (2.45)].
3	Вычисление производных таблично заданной функции	1. Вычислить первую и вторую производные во внутренних точках с помощью центральных разностей [см. формулы (3.18) и (3.20)]. 2. Вычислить первую и вторую производные в граничных точках с погрешностями $O(h)$ и $O(h^2)$ [см. формулы (3.16), (3.27), (3.29), (3.31), (3.33)].
4	Решение задачи Коши для обыкновенного дифференциального уравнения второго порядка	1. Записать уравнение в виде системы двух уравнений первого порядка. 2. Записать начальные условия для системы. 3. Найти решение системы, сделав два шага методом Эйлера или методом Эйлера с пересчетом [см. формулы (5.26), (5.29), (5.30)].

№ п/п	Умение	Алгоритмы
5	Решение краевой задачи для обыкновенного дифференциального уравнения второго порядка	1. Написать конечно-разностные уравнения для внутренних точек отрезка. 2. Написать конечно-разностную аппроксимацию для граничных условий. 3. Решить полученную систему методом Гаусса.

1. АППРОКСИМАЦИЯ И ИНТЕРПОЛЯЦИЯ ФУНКЦИЙ

1.1. Постановка задачи и основные определения

Цель рассматриваемых задач – найти аналитическое выражение (формулу) для описания функции $y = f(x)$. Далее эта формула может быть использована для исследования функции методами математического анализа (например, для вычисления производных, интегралов и т.д.).

Существуют два основных случая, когда на практике необходимо получение формулы для написания функции.

Первый и наиболее часто встречающийся случай заключается в том, что функция $y = f(x)$ задана таблично. Это означает, что дискретному множеству значений аргумента $X\{x_0, x_1, \dots, x_n\}$ поставлено в соответствие множество значений функции $Y\{y_0, y_1, \dots, y_n\}$. Численные значения x_i и y_i могут быть определены либо в результате эксперимента, либо путем численных расчетов. В этом случае аналитическое описание функции дает единственный способ вычисления функции $y = f(x)$ при любых значениях $x \neq x_i$ и для вычисления производных и интегралов от этой функции.

Второй случай заключается в том, что аналитическая зависимость $y = f(x)$ известна, но имеет очень громоздкий вид (например, в описание $f(x)$ входят очень сложные и громоздкие функции, вычисление которых на компьютере занимает много времени). В этом случае имеет смысл найти более простое аналитическое описание функции $y = f(x)$, которое позволит значительно уменьшить время ее вычисления. При этом, конечно, необходимо потребовать, чтобы полученная более простая функция была достаточно близкой к исходной. Понятие близости этих функций должно быть определено в зависимости от типа конкретной практической задачи.

Сформулируем **задачу о приближении (аппроксимации) функций**. Пусть задана функция $f(x)$ и требуется приближенно заменить (аппроксимировать) функцию $f(x)$ с помощью другой функции $\varphi(x)$ таким образом, чтобы отклонение $\varphi(x)$ от $f(x)$ (в некотором определенном смысле) было наименьшим в заданной области. Обычно функцию $f(x)$ называют аппроксимируемой, а функцию $\varphi(x)$ – аппроксимирующей.

* Жирным шрифтом выделены новые понятия, которые необходимо усвоить. Знание этих понятий будет проверяться при тестировании.

Аппроксимация может быть точечной или непрерывной. Аппроксимация называется **точечной**, если аппроксимирующая функция строится на дискретном множестве точек $X\{x_0, x_1, \dots, x_n\}$. Аппроксимация называется непрерывной (или интегральной), если $\varphi(x)$ строится на непрерывном множестве точек (например, на отрезке $[a, b]$).

При постановке задачи аппроксимации должны быть решены следующие вопросы:

1. *Как выбирать узловые точки $x_i \in X$, на базе которых строится аппроксимация, если имеется некоторая свобода в их выборе? В некоторых случаях удобно использовать равноотстоящие значения аргумента, т.е. полагать $x_i - x_{i-1} = h = \text{const}$ для всех значений i . Но иногда за счет специального распределения узлов x_i аппроксимируемой функции $f(x)$ можно уменьшить погрешность аппроксимации и даже добиться сходимости в том случае, когда она отсутствует при равномерном распределении узловых точек.*

2. *Какой класс функций выбрать при построении аппроксимирующей функции $\varphi(x)$? На практике чаще всего в качестве аппроксимирующей функции используются многочлены вида*

$$\varphi(x) = a_0 + a_1x + a_2x^2 + \dots + a_mx^m. \quad (1.1)$$

Такие приближения называют **многочленными приближениями**. Коэффициенты многочлена a_i и, быть может, степень многочлена m подбираются таким образом, чтобы обеспечить наименьшее отклонение $\varphi(x)$ от $f(x)$. Вопрос о том, какой смысл вкладывается в понятие наименьшее отклонение, зависит от конкретного вида решаемой задачи и будет разобран ниже. Удобство использования многочленов является очевидным. Во-первых, значения многочленов легко вычислять (например, с помощью схемы Горнера). Во-вторых, многочлены легко складывать, умножать, дифференцировать и интегрировать.

Другой класс аппроксимирующих функций – это тригонометрические многочлены

$$\varphi(x) = a_0/2 + \sum_{k=1}^m (a_k \cos(kx) + b_k \sin(kx)). \quad (1.2)$$

Приближения вида (1.2) называют рядами Фурье, теория которых достаточно хорошо разработана и которые широко используются в практических задачах.

3. *Каким должен быть критерий близости аппроксимируемой функции $f(x)$ и аппроксимирующей ее функции $\varphi(x)$?* Выбор критерия определяется существом решаемой задачи и тем, насколько точно задана исходная информация о функции $f(x)$ (например, тем, точно или приближенно известна таблица значений аргумента и функции $f(x)$ при табличном задании этой функции) и как далее нужно использовать полученную аппроксимирующую функцию $\varphi(x)$. Наиболее часто используются следующие критерии близости функций $f(x)$ и $\varphi(x)$:

а) *совпадение функций $f(x)$ и $\varphi(x)$ на заданном дискретном множестве точек x_i ($i = 0, 1, \dots, n$), на котором задана исходная функция $f(x)$*

$$\varphi(x_i) = f(x_i), \quad i = 0, 1, 2, \dots, n. \quad (1.3)$$

Такой тип точечной аппроксимации называется **интерполированием**. Обычно точки x_i называют **узлами интерполяции**, а функцию $\varphi(x)$, если она ищется в виде многочлена, – **интерполяционным многочленом**. Таким образом, близость интерполяционного многочлена $\varphi(x)$ к исходной функции $f(x)$ состоит в том, что их значения совпадают на заданном дискретном множестве точек x_i , но могут отличаться при $x \neq x_i$, где $x \in [x_0, x_n]$. Если интерполяционный многочлен ищется в виде (1.1) на всем интервале $[x_0, x_n]$, на котором задана функция $f(x)$, то $m = n$ и коэффициенты многочлена $a_0, a_1, \dots, a_n$ в формуле (1.1) ищутся из условий (1.3), которые в совокупности образуют систему линейных уравнений для величин a_i . Недостаток интерполяции проявляется по крайней мере в двух случаях: при большом количестве узлов интерполяции, когда степень многочлена $\varphi(x)$ становится большой; в том случае, когда табличные значения функции $f(x)$ получены экспериментальным путем или путем не очень точных вычислений и содержат в себе ошибки. В последнем случае не имеет смысла требовать обязательного прохождения аппроксимирующего многочлена $\varphi(x)$ через узловые точки x_i , $y_i = f(x_i)$ и удобнее применять другие критерии близости $f(x)$ и $\varphi(x)$, один из которых приведен ниже;

б) *критерий близости $f(x)$ и $\varphi(x)$, определяемый среднеквадратичным приближением*, в котором мерой близости функций $f(x)$ и $\varphi(x)$ является величина S , определяемая формулой

$$S = \sum_{i=0}^n [\varphi(x_i) - y_i]^2. \quad (1.4)$$

Величина S представляет собой сумму квадратов отклонений аппроксимирующей функции $\varphi(x)$ от исходной функции $f(x)$ на заданной системе точек $x_0, x_1, \dots, x_n$. Для заданной функции $f(x)$ естественно выбирать $\varphi(x)$ так, чтобы величина S , определяемая по (1.4), была минимальна. В случае многочленного приближения, когда $\varphi(x)$ берется в виде (1.1), коэффициенты многочлена $a_0, a_1, \dots, a_m$ подбирают из условия, что величина S является минимальной. Такой подход определяет сущность аппроксимации, даваемой **методом наименьших квадратов**. Подчеркнем, что в этом методе не требуется, чтобы функция $\varphi(x)$ проходила через узлы, в которых задана исходная табличная функция $f(x)$, а близость функций $\varphi(x)$ и $f(x)$ понимается лишь "в среднем", в смысле, даваемом формулой (1.4). При этом, конечно, в отдельных точках функции $\varphi(x)$ и $f(x)$ могут отличаться довольно значительно. Впрочем, такой же недостаток имеет и метод интерполяции, в котором в точках, расположенных между узлами интерполяции, функции $\varphi(x)$ и $f(x)$ могут значительно отличаться.

Однако в некоторых практических задачах и теоретических исследованиях требуется, чтобы во всех точках отрезка $[a, b]$, на котором задана функция $f(x)$, отклонение многочлена $\varphi(x)$ от функции $f(x)$ было по абсолютной величине меньше, чем заданная малая величина $\varepsilon > 0$:

$$|\varphi(x) - f(x)| < \varepsilon \quad (x \in [a, b]). \quad (1.5)$$

В этом случае говорят, что $\varphi(x)$ равномерно аппроксимирует функцию $f(x)$ на отрезке $[a, b]$. Критерий близости функций $\varphi(x)$ и $f(x)$ здесь будет уже другим;

в) *критерий близости функций при равномерном приближении* определяется понятием **абсолютного отклонения Δ функций $\varphi(x)$ и $f(x)$**

$$\Delta = \max_{a \leq x \leq b} |\varphi(x) - f(x)|. \quad (1.6)$$

Условие равномерной аппроксимации функции $f(x)$ с помощью $\varphi(x)$ с точностью ε означает, что $\Delta < \varepsilon$. В этом случае гарантируется, что

при любых $x \in [a, b]$ разность между значениями функций $\varphi(x)$ и $f(x)$ не превышает ε :

$$\max_{a \leq x \leq b} |\varphi(x) - f(x)| < \varepsilon. \quad (1.7)$$

Возможность построения многочлена, равномерно приближающего функцию $f(x)$, определяется теоремой Вейерштрасса об аппроксимации.

Теорема. Если функция $f(x)$ непрерывна на отрезке $[a, b]$, то для любого $\varepsilon > 0$ существует многочлен $\varphi(x)$, степень которого зависит от ε , $m = m(\varepsilon)$, абсолютное отклонение которого от функции $f(x)$ на отрезке $[a, b]$ меньше ε .

Различия трех рассмотренных критериев приближения функций – интерполяции, среднеквадратичного приближения и равномерного приближения – приведены на рис 1.1, где а – интерполяция; б – среднеквадратичное приближение; в – равномерное приближение.

Из теоремы Вейерштрасса следует, что при заданной точности аппроксимации ε можно подобрать соответствующий многочлен степени m , равномерно приближающий заданную функцию $f(x)$. Однако наряду с выбором m возможен и другой подход, при котором степень аппроксимирующего полинома m является фиксированной и требуется подобрать такой многочлен степени m , чтобы на заданном отрезке $[a, b]$ величина абсолютного отклонения Δ , определяемая по (1.6), была минимальна

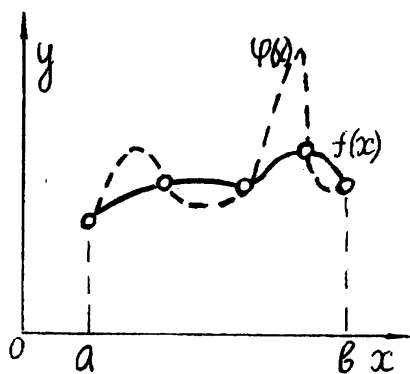
$$\Delta = \Delta_{\min}. \quad (1.8)$$

Такой многочлен $\varphi(x)$ называется **многочленом наилучшего равномерного приближения**. Выполнение (1.8) осуществляется за счет подбора коэффициентов $a_0, a_1, \dots, a_m$ многочлена (1.1). Можно показать, что для непрерывной на замкнутом ограниченном множестве функции $f(x)$ такой многочлен $\varphi(x)$ существует и является единственным.

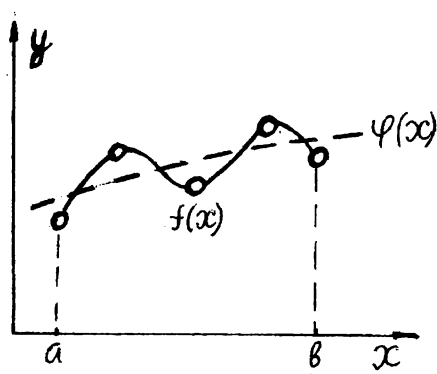
Соотношение (1.8) можно записать в виде

$$\Delta = \min_{\varphi(x)} \left(\max_{a \leq x \leq b} |\varphi(x) - f(x)| \right). \quad (1.9)$$

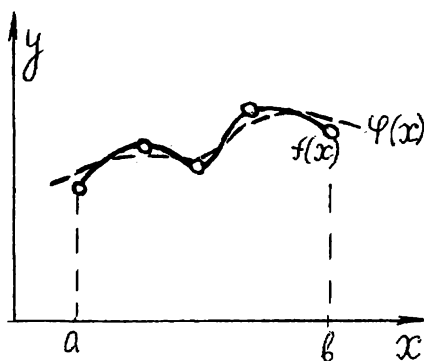
Задачи такого типа часто называют минимаксными. В их разработку большой вклад внес знаменитый русский математик П. Л. Чебышев.



а)



б)



в)

Рис. 1.1

Смысл соотношения (1.9) заключается в том, что мы подбираем такой многочлен степени m ($\varphi(x) = P_m(x)$), который на отрезке $[a, b]$ дает наименьшую абсолютную ошибку.

1.2. Интерполяция с помощью многочленов

Рассмотрим многочленную интерполяцию функции $f(x)$ с помощью многочлена $\varphi(x)$

$$\varphi(x) = a_0 + a_1x + a_2x^2 + \dots + a_mx^m. \quad (1.10)$$

Будем считать, что функция $f(x)$ задана таблично и, следовательно, задано множество значений ее аргумента $\{x_0, x_1, \dots, x_n\}$ и соответствующее множество значений функции $\{y_0, y_1, \dots, y_n\}$, где $y_i = f(x_i)$.

Различают два вида интерполяции: глобальную и локальную. Интерполяция называется **глобальной**, если один многочлен (1.10) используется для интерполяции функции $f(x)$ на всем рассматриваемом интервале изменения ее аргумента x . В этом случае $m = n$ и коэффициенты интерполяционного многочлена $a_0, a_1, \dots, a_n$ определяются из решения системы линейных уравнений (1.3), которая в данном случае имеет вид

[illegible]

Интерполяция называется **локальной** (или кусочной), если интерполяционные многочлены строятся отдельно для разных частей рассматриваемого интервала изменения x . В этом случае $m < n$. Идея такого подхода основана на том, чтобы многочленами невысокой степени интерполировать функцию $f(x)$ на отдельных интервалах ее аргумента, подбирая для каждого интервала свой наилучший многочлен.

1.2.1. Интерполяционный многочлен Лагранжа

Рассмотрим случай глобальной интерполяции с произвольным расположением узлов интерполяции x_i ($i = 0, 1, \dots, n$). Для нахождения коэффициентов полинома (1.10) необходимо решать систему линейных уравнений (1.11). Такой путь является достаточно громоздким и поэтому разработаны более простые алгоритмы получения интерполяционных многочленов, один из которых принадлежит Лагранжу. Рассмотрим выражение

$$L(x) = \sum_{i=0}^n y_i \frac{(x-x_0)(x-x_1)\dots(x-x_{i-1})(x-x_{i+1})\dots(x-x_n)}{(x_i-x_0)(x_i-x_1)\dots(x_i-x_{i-1})(x_i-x_{i+1})\dots(x_i-x_n)} \quad (1.12)$$

$L(x)$ представляет собой многочлен степени n от переменной x , в чем можно убедиться, перемножив двучлены, стоящие в числителе каждого члена суммы, и учитывая, что величины x_i и y_i являются заданными числами, определяющими таблично заданную функцию $y = f(x)$. Структура многочлена (1.12) такова, что он удовлетворяет основным условиям интерполяции: значения этого многочлена при $x = x_i$ равны значениям интерполируемой функции $y = f(x)$ в этих точках $y_i = f(x_i)$. Действительно, дробь, на которую умножается y_i в каждом члене суммы, равна единице при $x = x_i$ (так как числитель дроби равен ее знаменателю) и нулю при $x = x_j$ ($j \neq i$) (так как при $x = x_j$ числитель дроби обращается в нуль). Следовательно,

$$L(x_i) = y_i \quad (i = 0, 1, \dots, n) \quad (1.13)$$

и многочлен $L(x)$ является интерполяционным многочленом для функции $y = f(x)$. Многочлен $L(x)$ называют интерполяционным многочленом Лагранжа. Можно показать, что для заданной функции $f(x)$ этот многочлен является единственным. С помощью $L(x)$ можно определить значения функции $f(x)$ при любых значениях x , не являющихся узлами x_i функции $f(x)$. Из формулы (1.12), как частные случаи, можно получить формулы для **линейной и квадратичной интерполяции**, которые часто используются на практике.

Линейная интерполяция – простейший случай интерполяции многочленами, когда интерполирующей функцией является кусочно-линейная функция (ломаная).

При $n = 1$

$$L(x) = \frac{x - x_1}{x_0 - x_1} y_0 + \frac{x - x_0}{x_1 - x_0} y_1. \quad (1.14)$$

Легко видеть, что $L(x_0) = y_0$, $L(x_1) = y_1$. Выражение (1.14) дает формулу линейной интерполяции для приближенного определения значения $f(x)$ при любом значении x , лежащем внутри интервала $[x_0, x_1]$, на концах которого значения функции известны: $y_0 = f(x_0)$, $y_1 = f(x_1)$.

При $n = 2$

$$L(x) = \frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)} y_0 + \frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)} y_1 + \frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)} y_2. \quad (1.15)$$

Видно, что $L(x_0) = y_0$, $L(x_1) = y_1$, $L(x_2) = y_2$. Квадратный многочлен (1.15) дает возможность по значениям y_0, y_1, y_2 функции $f(x)$ в трех заданных точках x_0, x_1, x_2 определить приближенное значение $f(x)$ в любой точке $x \in [x_0, x_2]$. Формула (1.15) называется формулой квадратичной интерполяции.

Приведем примеры использования формул (1.14) и (1.15).

Пример 1. Функция $y = f(x)$ задана таблицей:

x	1	2	3
y	0	0,693147	1,0986123

Вычислить значение функции $f(x)$ в точке $x = 1,5$, используя линейную и квадратичную интерполяцию.

Решение. В нашем случае $x_0 = 1$, $x_1 = 2$, $x_2 = 3$, $y_0 = 0$; $y_1 = 0,69315$; $y_2 = 1,09861$. Для вычисления $f(1,5)$ воспользуемся формулой линейной интерполяции (1.14)

$$f(1,5) = \frac{1,5-1}{-1} \cdot 0 + \frac{1,5-1}{+1} \cdot 0,69315 = -0,5 \cdot 0 + 0,5 \cdot 0,69315 = +0,346574.$$

Воспользуемся теперь формулой квадратичной интерполяции (1.15)

$$\begin{aligned} f(1,5) &= \frac{(1,5-2)(1,5-3)}{(1-2)(1-3)} y_0 + \frac{(1,5-1)(1,5-3)}{(2-1)(2-3)} y_1 + \frac{(1,5-1)(1,5-2)}{(3-1)(2-1)} y_2 = \\ &= \frac{(-0,5) \cdot (-1,5)}{2} \cdot 0 + \frac{0,5 \cdot (-1,5)}{-1} \cdot 0,693147 + \frac{0,5 \cdot (-0,5)}{2} \cdot 1,098612 = \\ &= 0,519863 - 0,137327 = 0,382536. \end{aligned}$$

Для сравнения результатов укажем, что приведенная в таблице функция – это функция $y = \ln x$, в которой были использованы шесть

знаков после запятой. Точное значение $y(1,5)$ будет равно $\ln(1,5) = 0,405465$, линейная интерполяция дает $y(1,5) = 0,346574$, а квадратичная интерполяция – $y(1,5)=0,382536$. Для линейной интерполяции ошибка равна 0,058891, а для квадратичной – 0,022929.

1.2.2. Интерполяционный многочлен Ньютона

В многочлене Лагранжа (1.12) используются произвольно расположенные узлы интерполяции x_i . Рассмотрим теперь часто встречающийся случай равноотстоящих узлов, когда $x_i - x_{i-1} = h = \text{const}$ ($i = 1, 2, \dots, n$). Величина h называется шагом. Получим для этого случая другой способ записи интерполяционного многочлена, называемый интерполяционным многочленом Ньютона. Подчеркнем здесь, что при заданном наборе узлов интерполяции интерполяционный многочлен является единственным. Поэтому различные виды интерполяционных многочленов (Лагранжа, Ньютона и многие другие) различаются лишь алгоритмами их получения и способом записи, представляя собой в сущности один и тот же многочлен. Удобство применения конкретного вида записи интерполяционного многочлена определяется существом решаемой задачи.

Для дальнейшего нам понадобится ввести понятие **конечных разностей**. В следующих разделах это понятие будет широко использоваться при изложении численных методов решения дифференциальных уравнений. Пусть известны значения функции $y = f(x)$ в узлах x_i : $y_i = f(x_i)$. Составим разности значений функции в соседних узлах:

$$\begin{array}{llll} \Delta y_0 & = y_1 - y_0 & = f(x_0 + h) - f(x_0) \\ \Delta y_1 & = y_2 - y_1 & = f(x_0 + 2h) - f(x_0 + h) \\ \dots & \dots & \dots & \dots \\ \Delta y_{n-1} & = y_n - y_{n-1} & = f(x_0 + nh) - f(x_0 + (n-1)h) \end{array} \quad (1.16)$$

Величины Δy_i называются **первыми разностями функции** $y = f(x)$. Имея первые разности, можно составить **вторые разности функции**:

Многочлен (1.20) позволяет определить значения функции $y = f(x)$ при любом значении $x \neq x_i$ на всем интервале $[x_0, x_n]$. [Напомним, что в узлах интерполяции $y_i = f(x_i) = N(x_i)$]. Для этого по величине x достаточно определить значение $t = (x - x_0)/h$ и воспользоваться формулой (1.20), определив предварительно конечные разности функции $y = f(x)$ по формулам (1.16)–(1.18). Однако чем меньше значение x при $x \in [x_0, x_n]$, тем меньше членов в сумме (1.20) приходится брать (значения слагаемых быстро убывают) и тем точнее получается результат интерполяции. Поэтому (1.20) целесообразно использовать при $x_0 < x < x_1$ (и, следовательно, при $t < 1$). Для значений аргумента, например, при $x_1 < x < x_2$ вместо x_0 удобнее взять x_1 , т.е. левую точку соответствующего интервала. В общем случае, если $x \in [x_i, x_{i+1}]$ и в качестве базисной точки, в которой считаются конечные разности, берется значение x_i , интерполяционный многочлен Ньютона можно записать в виде более общем, чем (1.20)

$$N(x_i + th) = y_i + t\Delta y_i + \frac{t(t-1)}{2!}\Delta^2 y_i + \dots + \frac{t(t-1)\dots(t-n+1)}{n!}\Delta^n y_i \quad (1.21)$$

($i = 0, 1, \dots$).

Выражение (1.21) называется интерполяционным многочленом Ньютона для интерполирования вперед. Удобство использования (1.21) заключается в том, что беря в качестве базисной точки x_i левый узел интерполяции ($x \in [x_i, x_{i+1}]$), мы получаем ряд (1.21), в котором первое слагаемое будет главным членом, а остальные слагаемые будут носить характер поправок, причем значения поправок будут быстро уменьшаться для более далеких членов ряда.

В частном случае из формул (1.19), (1.20) или (1.21) можно получить формулы для линейной или квадратичной интерполяции в случае равноотстоящих узлов. Из (1.19) при $n = 1$, когда $x \in [x_0, x_1]$, получим формулу линейной интерполяции

$$y \approx y_0 + \frac{y_1 - y_0}{h}(x - x_0), \quad (1.22)$$

а при $n = 2$, когда $x \in [x_0, x_2]$, получим формулу квадратичной интерполяции

$$y \approx y_0 + \frac{y_1 - y_0}{h}(x - x_0) + \frac{y_2 - 2y_1 + y_0}{2h^2}(x - x_0)(x - x_1). \quad (1.23)$$

1.2.3. Точность и сходимость многочленной интерполяции

По определению интерполяции интерполяционные многочлены $L(x)$ и $N(x)$ совпадают в узлах интерполяции x_i со значениями функции $y = f(x)$. Однако при других $x \neq x_i$ значения этих многочленов и функции $f(x)$ не совпадают. Без вывода приведем формулы для оценки погрешности аппроксимации интерполяционных многочленов Лагранжа и Ньютона. Обозначим через $L_n(x)$ и $N_n(x)$ многочлены Лагранжа и Ньютона n -ой степени. Рассмотрим погрешности интерполяции:

$$R_n(x) = f(x) - L_n(x); \quad (1.24)$$

$$Z_n(x) = f(x) - N_n(x). \quad (1.25)$$

Иногда величины $R_n(x)$ и $Z_n(x)$ называют остаточными членами многочленов Лагранжа и Ньютона. Можно показать, что для этих величин существуют следующие соотношения:

$$|R_n(x)| \leq \frac{(x - x_0)(x - x_1) \dots (x - x_n)}{(n + 1)!} M_{n+1}, \quad (1.26)$$

где $M_{n+1} = \max_{x_0 \leq x \leq x_n} |f^{(n+1)}(x)|$ – максимальное значение производной порядка $(n + 1)$ функции $y = f(x)$ на отрезке $[x_0, x_n]$;

$$Z_n(x) = \frac{t(t-1) \dots (t-n)}{(n+1)!} f^{(n+1)}(\zeta) h^{n+1}, \quad (1.27)$$

где $\zeta \in [x_0, x_n]$, а $t = (x - x_0)/h$.

Из выражений (1.26) и (1.27) видно, что ошибка интерполяции зависит, в частности, от величины $(n + 1)$ -ой производной интерполи-

руемой функции $f(x)$. Часто в математике используется понятие “гладкости” функции. Говоря о поведении функции, имеют в виду улучшение или ухудшение ее гладкости, т.е. поведение ее n -ой производной $f^{(n)}(x)$ при увеличении n . Не давая точных определений этого понятия, которые будут рассмотрены в функциональном анализе, будем считать, что рост величины $f^{(n)}(x)$ при увеличении n означает ухудшение гладкости функции, а уменьшение величины производных с ростом n – улучшение ее гладкости. Типичным примером негладкой функции является функция $y = \sin(5x)$, для которой

$$y' = 5 \cos(5x), \quad y'' = -5^2 \sin(5x), \quad y''' = -5^3 \cos(5x), \quad y^{IV} = 5^4 \sin(5x)$$

и т.д. Видно, что $\max_x |y^{(n)}(x)| = 5^n \rightarrow \infty$ при $n \rightarrow \infty$.

Оценка погрешности интерполяционных многочленов показывает, что с ростом числа узлов n и соответственно степени многочлена $P_n(x)$ погрешность может увеличиваться из-за роста $|f^{(n)}(x)|$. Кроме того, с ростом n увеличивается время вычисления многочлена на ЭВМ и возрастает вычислительная погрешность. Это является одним из главных недостатков глобальной интерполяции, когда ищется единый интерполяционный многочлен на всем интервале $[x_0, x_n]$, где задана функция $f(x)$. Возможны два способа преодоления этих трудностей. Первый способ основан на использовании локальной интерполяции, когда на каждом интервале $[x_i, x_{i+1}]$ строится свой многочлен невысокой степени (обычно $m = 1, 2$ или 3). Соответствующие сплайновые интерполяции описываются в подразделе 1.2.4.

Второй способ основан на специальном расположении узлов интерполяции x_i (если они не фиксированы) и требует некоторых пояснений. Было установлено, что многочленная аппроксимация с равномерно расположенными узлами интерполяции x_i может приводить к совершенно неверным результатам при увеличении степени n интерполяционного многочлена. Впервые такой пример был приведен О. Рунге, который интерполировал функцию

$$y = \frac{1}{1 + 25x^2}$$

на интервале $[-1, 1]$ с равномерным распределением узлов

интерполяции. Оказалось, что при любых x , для которых $|x| < 0,726$, интерполяционный многочлен быстро сходится к соответствующему значению $f(x)$ при $n \rightarrow \infty$, а на концах интервала при $|x| > 0,726$ интерполяционный многочлен не стремится ни к какому пределу и возрастает бесконечно в каждой точке этой части интервала при $n \rightarrow \infty$. Такое поведение интерполяционного многочлена при равномерном распределении узлов, когда при $n \rightarrow \infty$ значения этого многочлена на одной части интервала сходятся к $f(x)$, а на другой – не сходятся, принято называть явлением Рунге. Было показано, что таких затруднений не возникает, если специальным образом выбирать неравномерно расположенные узлы интерполяции. В частности, доказано, что если функция $f(x)$ имеет непрерывную производную на отрезке $[-1, 1]$, то при выборе узлов интерполяции x_i , совпадающих с корнями многочленов Чебышева степени $n + 1$, интерполяционные многочлены степени n сходятся к значениям $f(x)$ в любой точке этого отрезка. Связанные с этим вопросы подробно разобраны в работе [7].

1.2.4. Использование локальных интерполяций

В силу отмеченных выше недостатков глобальной интерполяции одним многочленом на всем отрезке $[x_0, x_n]$, на практике часто используется локальная интерполяция, которую можно назвать кусочно-многочленной. Идея ее заключается в том, чтобы на каждом отдельном отрезке $[x_i, x_{i+1}]$ использовать свой многочлен степени m ($m < n$), коэффициенты которого определяются из условий совпадения значений этого многочлена в узловых точках со значениями функции $f(x)$ в тех же точках и из некоторых условий сопряжения локальных многочленов в узловых точках с тем, чтобы обеспечить необходимую гладкость интерполирующей функции. Простейшим случаем локальной интерполяции является кусочно-линейная интерполяция, когда между узлами интерполяции функция ведет себя линейно (рис. 1.2). В этом случае на каждом интервале $[x_i, x_{i+1}]$ используется уже рассмотренная ранее формула линейной интерполяции (1.14).

Наиболее часто используются кубические сплайны, в которых на каждом интервале $[x_i, x_{i+1}]$ для интерполирования используется свой кубический многочлен, причем многочлены на соседних интервалах стыкуются определенным образом в узлах интерполяции x_i . Под термином сплайн подразумевается следующее понятие. Сплайном

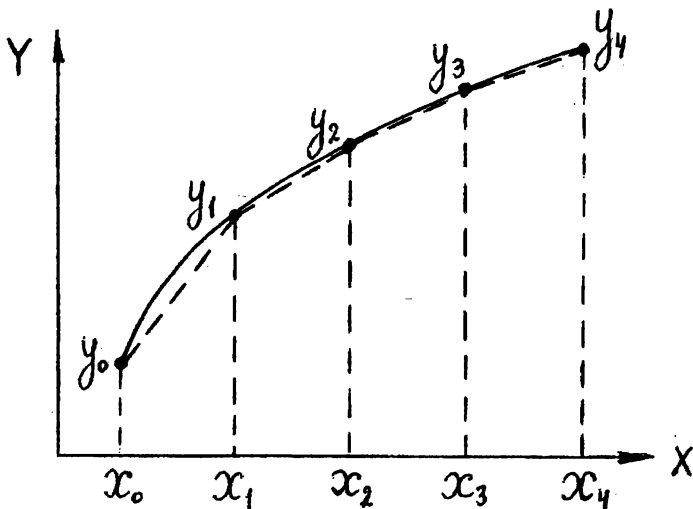


Рис. 1.2. Кусочно-линейная интерполяция

порядка m называют функцию $S_m(x)$, являющуюся многочленом порядка m на каждом из отрезков $[x_i, x_{i+1}]$

$$S_m(x) = P_m(x) = a_{i0} + a_{i1}x + \dots + a_{im}x^m \quad (x \in [x_i, x_{i+1}]) \quad (1.28)$$

и удовлетворяющую условиям непрерывности производных до порядка $m - 1$ во внутренних узлах интерполяции $x_1, x_2, \dots, x_{n-1}$:

$$P_{mi}^{(k)}(x_i) = P_{m,i+1}^{(k)}(x_i) \text{ при } k = 0, 1, \dots, m-1 \quad (i = 1, 2, \dots, n-1), \quad (1.29)$$

где верхний индекс k означает порядок производной.

Кубический сплайн $S_3(x)$ имеет вид

$$S_3(x) = P_{3i}(x) = a_{i0} + a_{i1}x + a_{i2}x^2 + a_{i3}x^3,$$

где коэффициенты кубического многочлена будут различными для каждого интервала $[x_i, x_{i+1}]$. Так как на общем интервале $[x_0, x_n]$ всего n интервалов $[x_i, x_{i+1}]$, то общее количество неизвестных коэффици-

ентов у многочленов равно $4n$. Для определения $4n$ неизвестных могут быть написаны $2n$ уравнений, выражающих условия прохождения функции $S_3(x)$ через заданные точки $S_3(x_i) = y_i$, $S_3(x_{i+1}) = y_{i+1}$ и $2n - 2$ уравнения вида (1.29), выражающих условия непрерывности первых и вторых производных в узлах интерполяции (условия гладкого сопряжения многочленов на соседних интервалах). Два недостающих соотношения обычно задаются на концах интервала при $x = x_0$ и $x = x_n$ и выражают условия закрепления концов сплайна. В частности, при свободном закреплении концов полагают $S_3''(x_0) = 0$ и $S_3''(x_n) = 0$. Получаемая функция, называемая свободным кубическим сплайном, является самой гладкой среди всех интерполяционных функций данного класса. Имеются стандартные программы для различных алгоритмических языков, которые позволяют для исходной табличной функции $f(x)$ построить соответствующие сплайновые интерполяции и вычислить их значение при любом x ($x \in [x_0, x_n]$).

1.3. Равномерные многочленные приближения

1.3.1. Использование степенных разложений для вычисления функций

Одной из главных задач, которую приходится решать при работе на компьютере, является вычисление разнообразных функций, как элементарных, таких как $\sin x$, $\ln x$, e^x и т.п., так и специальных, которые используются в различных областях науки и техники. Обычно в любом алгоритмическом языке имеется набор стандартных подпрограмм, которые реализуют их вычисления. В частности, всегда имеется набор подпрограмм, вычисляющих элементарные функции. Так как они используются чрезвычайно часто, то важно разработать эффективные и экономичные алгоритмы, которые позволяют вычислять эти функции с высокой точностью за сравнительно небольшое количество операций. Аналогично обстоит дело и при вычислении любых других функций. В этом разделе будет приведен один из экономичных алгоритмов, использующий свертывание степенных разложений с помощью многочленов Чебышева. П.Л. Чебышев (1821 – 1894) – знаменитый русский математик, являющийся основоположником раздела математики, посвященного теории равномерного приближения функций. Разработанные им идеи положены в основу многих других разделов современной вычислительной математики.

Рассмотрим возможность вычисления элементарных функций на ЭВМ. При ручном счете для этой цели обычно используются таблицы,

которые содержат вычисленные значения функций с заданным (обычно постоянным) шагом изменения аргумента. При использовании ЭВМ такой способ неудобен в силу следующих причин. Во-первых, для ввода подобных таблиц потребуется большой объем оперативной памяти. Во-вторых, поиск нужного значения функции может потребовать выполнения большого количества операций. Если же вводить сравнительно небольшую таблицу (с большим шагом изменения аргумента), то для получения значения функции с высокой точностью будет недостаточно простой линейной или квадратичной интерполяции и придется использовать более сложные интерполяционные многочлены (например, Лагранжа или Ньютона), что приведет к увеличению времени вычисления функции. Выход состоит в том, чтобы для вычисления функций использовать экономичные степенные разложения, для получения которых требуется сравнительно немного арифметических операций. На первый взгляд, для этой цели можно использовать степенные ряды типа рядов Тейлора и Маклорена, которые ранее изучались в курсе высшей математики. Возникающие здесь трудности рассмотрим на примере вычисления функции $\sin x$ на отрезке $[-1, 1]$. Напомним, что разложение функции в ряд Тейлора на отрезке $[a, b]$ имеет вид

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \dots + \frac{f^{(n)}(a)}{n!}(x-a)^n + R_n, \quad (1.30)$$

где $x = a$ – базовая точка, относительно которой ищется разложение, а R_n – остаточный член, который определяет погрешность разложения в том случае, если мы ограничимся $(n+1)$ -ым членом разложения. Оценка этой погрешности имеет вид

$$\max_{a \leq x \leq b} |R_n| \leq \frac{|b-a|^{n+1}}{(n+1)!} \max_{a \leq x \leq b} |f^{(n+1)}(x)|. \quad (1.31)$$

Ряд Тейлора (1.30) выражает тот замечательный факт, что, зная значения функции $f(x)$ и ее производных лишь в одной точке $x = a$, мы можем вычислить значение этой функции на всем отрезке $[a, b]$, на котором сходится ряд (1.30).

Взяв за базовую точку $x = 0$, для функции $\sin x$ получим

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots; \quad (1.32)$$

Теоретически ряд (1.32) сходится при любом x . Однако практически

для его вычисления при достаточно больших x требуется брать очень много членов ряда, так как сходимость будет чрезвычайно медленной и, кроме того, будут сильно расти погрешности округления. Отчасти выйти из положения можно, воспользовавшись свойством периодичности функции $\sin x$ и ограничившись отрезком $[-\pi/2, \pi/2]$. Однако все равно разложение (1.32), как и любой другой ряд Тейлора, будет обладать тем недостатком, что вблизи базовой точки (в данном случае $x = 0$) $f(x)$ вычисляется с очень высокой точностью, а в остальной части интервала точность будет гораздо меньше и придется вычислять очень много членов ряда. Это – фундаментальный недостаток степенных разложений Тейлора, из-за чего они мало используются в практических вычислениях, и вместо них применяются гораздо более эффективные разложения, в частности, разложения с использованием многочленов Чебышева. В этом случае погрешность достаточно равномерно распределена на всем интервале, где вычисляется функция $f(x)$.

1.3.2. Многочлены Чебышева

Рассмотрим многочлены Чебышева, которые очень широко используются в вычислительной математике и обладают многими ценными свойствами. Многочлены Чебышева $T_n(x)$ степени n определяются формулой

$$T_n(x) = \frac{1}{2} \left[\left(x + \sqrt{x^2 - 1} \right)^n + \left(x - \sqrt{x^2 - 1} \right)^n \right] \quad (1.33)$$

$$(-1 \leq x \leq 1, \quad n = 0, 1, 2, \dots).$$

Можно показать, что (1.33) действительно представляют собой многочлены, так как члены, содержащие корни, при возведении в степень и последующих преобразованиях пропадают. Проведя соответствующие вычисления, получим:

$$T_0(x) = 1, \quad T_1(x) = x, \quad T_2(x) = 2x^2 - 1, \quad T_3(x) = 4x^3 - 3x,$$

$$T_4(x) = 8x^4 - 8x^2 + 1, \quad T_5(x) = 16x^5 - 20x^3 + 5x, \quad \dots \quad (1.34)$$

Очень удобным свойством является то, что коэффициенты многочлена – целые числа. Кроме того, для вычисления многочленов можно воспользоваться простым рекуррентным соотношением

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x) \quad (n = 1, 2, 3, \dots), \quad (1.35)$$

которое позволяет последовательно определять многочлены со все более высокими степенями, начиная с двух первых многочленов $T_0(x)$ и $T_1(x)$. Можно показать, что коэффициент при старшем члене определяется формулой $a_n = 2^{n-1}$.

Чрезвычайно важным свойством многочленов Чебышева является тот факт, что они могут быть выражены через элементарные тригонометрические функции. Они представляют собой фактически простые тригонометрические функции $\cos(n\theta)$, только выраженные через переменную $x = \cos\theta$. Поэтому $T_n(x)$ можно представить в виде

$$T_n(x) = \cos(n \arccos x) \quad n = 0, 1, 2, \dots \quad (1.36)$$

Это является наиболее важным свойством многочленов Чебышева, так как означает, что разложение функции $f(x)$ по многочленам $T_n(x)$ является просто другой интерпретацией разложения четной функции в ряд по косинусам, что позволяет переносить многие важные свойства рядов Фурье в область степенных рядов. С помощью (1.36) можно получить выражения (1.33) и (1.35). Корни (нули) многочленов Чебышева на отрезке $[-1, 1]$ определяются формулой

$$x_k = \cos \frac{2k-1}{2n} \pi \quad (k = 0, 1, 2, \dots, n).$$

Они расположены неравномерно на отрезке $[-1, 1]$, сгущаясь к его концам. Точки экстремума $T_n(x)$ определяются соотношениями

$$x_k = \cos(k\pi/n) \quad (k = 1, 2, \dots, n-1),$$

причем в точках максимума $T_n(x) = 1$, а в точках минимума $T_n(x) = -1$. Это очень важное свойство, означающее, что всегда выполняется соотношение

$$|T_n(x)| \leq 1 \quad (1.37)$$

при любых n и x , что очень удобно для проведения различных оценок погрешностей при использовании многочленов Чебышева.

Рассмотрим многочлен $\overline{T}_n(x)$, полученный путем деления

многочлена $T_n(x)$ на коэффициент при старшем члене (при x^n)
 $a_n = 2^{n-1}$

$$\overline{T}_n(x) = \frac{1}{2^{n-1}} T_n(x). \quad (1.38)$$

У $\overline{T}_n(x)$ коэффициент при старшем члене будет равен 1. Наряду с многочленом $\overline{T}_n(x)$ рассмотрим любые другие многочлены $P_n(x)$ степени n , у которых коэффициент при старшем члене равен единице [как и у $\overline{T}_n(x)$]. Тогда можно доказать, что у многочлена $\overline{T}_n(x)$ максимальное абсолютное отклонение на интервале $[-1, 1]$ будет наименьшим по сравнению с любым другим многочленом $P_n(x)$. Это означает, что

$$\max_{-1 \leq x \leq 1} |P_n(x)| \geq \max_{-1 \leq x \leq 1} |\overline{T}_n(x)| = \frac{1}{2^{n-1}}. \quad (1.39)$$

Поэтому говорят, что многочлены Чебышева $\overline{T}_n(x)$ являются наименее уклоняющимися от нуля на интервале $[-1, 1]$. Выражение (1.39) иногда называют принципом минимакса. Это означает, что для $\overline{T}_n(x)$ наибольшее значение на $[-1, 1]$ будет наименьшим по сравнению с наибольшими значениями любых других многочленов $P_n(x)$.

В подразд. 1.1 рассматривались три возможных критерия согласия аппроксимируемой функции $f(x)$ и аппроксимирующей ее функции: совпадение в узловых точках (интерполяция); минимум среднеквадратичного отклонения; равномерное приближение, когда минимизируется абсолютное отклонение Δ функций $f(x)$ и $\varphi(x)$

$$\min \Delta = \min_{-1 \leq x \leq 1} \max |\varphi(x) - f(x)|, \quad (1.39a)$$

что является другой формой записи (1.39). Поэтому часто равномерное приближение называют "чебышевским приближением", при котором стремятся свести к минимуму максимальную величину ошибки $|f(x) - \varphi(x)|$ на всем интервале.

Подчеркнем разницу между среднеквадратичным и чебышевским приближениями. Среднеквадратичное приближение уменьшает среднюю квадратичную ошибку, но при этом может допускать большие ошибки в отдельных точках интервала. Чебышевское приближение

уменьшает наибольшую ошибку, допуская большое среднеквадратичное отклонение.

В заключение приведем формулы для выражения различных степеней x с помощью многочленов Чебышева. Для этого из формул (1.34) для многочленов Чебышева последовательно будем выражать все более высокие степени x . В результате получим:

$$\begin{aligned} 1 = T_0, \quad x = T_1, \quad x^2 = \frac{1}{2}(T_0 + T_2), \quad x^3 = \frac{1}{4}(3T_1 + T_3), \\ x^4 = \frac{1}{8}(3T_0 + 4T_2 + T_4), \quad x^5 = \frac{1}{16}(10T_1 + 5T_3 + T_5), \quad \dots \end{aligned} \quad (1.396)$$

1.3.3. Экономизация степенных рядов с помощью многочленов Чебышева

Одно из важных приложений многочленов Чебышева – их использование для эффективного вычисления степенных рядов, основанное на свойстве наилучшей сходимости этих многочленов по сравнению с любыми другими многочленами. Суть дела заключается в следующем. Одна и та же функция $f(x)$ может быть разложена в различные степенные ряды в зависимости от способа их построения. Эти ряды представляют одну и ту же функцию, но обладают различной скоростью сходимости. С теоретической точки зрения эти разложения равнозначны, так как в пределе будет получена одна и та же функция. Однако с практической точки зрения время вычисления функции на ЭВМ с помощью этих рядов может иногда отличаться в десятки и сотни раз. Пусть, например, нам известно разложение какой-либо функции в ряд Тейлора с конечным числом членов

$$f(x) = b_0 + b_1x + b_2x^2 + \dots + b_mx^m. \quad (1.40)$$

Пусть, с другой стороны, нам известно разложение этой же функции в ряд по многочленам Чебышева

$$f(x) = C_0 + C_1T_1(x) + C_2T_2(x) + \dots + C_mT_m(x). \quad (1.41)$$

Ряд Тейлора (1.40) часто имеет плохую сходимость почти всюду, кроме окрестности базовой точки, где заданы все значения производных. Ряд (1.41), напротив, имеет наилучшую сходимость. Это означает, что если оборвать разложение на члене m , как это сделано в

(1.40) и (1.41), то погрешность ряда (1.41) будет наименьшей и равномерно распределенной на всем интервале. Это дает возможность получить большую точность даже с малым числом членов.

Непосредственное разложение функции $f(x)$ в ряд по многочленам $T_n(x)$ в виде (1.41) удастся получить редко. Но существуют другие способы получения разложения вида (1.41). Они основаны на том, что мы исходим из ряда Тейлора (1.40) и с помощью специальных приемов сворачиваем его в более короткий ряд, не теряя существенно в точности. Подробное описание этих приемов, в частности эффективного τ -метода Ланцоша, приведено в работе [7]. На примере разберем один из таких простых приемов свертки ряда Тейлора с помощью чебышевских многочленов.

Пример 2. Получить наиболее эффективное степенное разложение для функции $y = \ln(1+x)$ в виде квадратного многочлена.

Решение. Степенной ряд Тейлора для этой функции имеет вид

$$y = \ln(1+x) \approx x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \frac{x^5}{5} + R_6, \quad (1.42)$$

где записано пять членов разложения, а остаточный член ряда Тейлора R_6 оценивается формулой (1.31). Однако в данном случае не требуется оценка для шестой производной, если учесть следующий факт из общей сходимости рядов. Так как ряд (1.42) является знакопеременным, то погрешность от замены бесконечного ряда его частичной суммой по модулю меньше первого отброшенного члена

$$|R_6| < \frac{x^6}{6} \leq \frac{1}{6} \quad (\text{при } 0 \leq x \leq 1). \quad (1.43)$$

Заменив теперь во всех членах разложения (1.42) соответствующие члены x^k с помощью формул (1.39б), получим разложение функции $y = \ln(1+x)$ в ряд по чебышевским многочленам

$$\begin{aligned} y = \ln(x) &= T_1 - \frac{1}{4}(T_0 + T_2) + \frac{1}{12}(3T_1 + T_3) - \\ &- \frac{1}{32}(3T_0 + 4T_2 + T_4) + \frac{1}{80}(10T_1 + 5T_3 + T_5) = \\ &= -\frac{11}{32}T_0 + \frac{11}{8}T_1(x) - \frac{3}{8}T_2(x) + \frac{7}{48}T_3(x) - \frac{1}{32}T_4(x) + \frac{1}{80}T_5(x). \end{aligned} \quad (1.44)$$

Ряд Тейлора (1.42) дает очень высокую точность в малой окрестности точки $x = 0$ и большую погрешность при больших значениях x в интервале $[0, 1]$, где этот ряд сходится. Согласно (1.43), наибольшая погрешность будет при $x = 1$: $|R_6| < \frac{1}{6}$. У ряда (1.44) погрешность будет распределена более равномерно и при $x = 1$ она будет существенно меньше, чем у ряда Тейлора (1.42). Но использование чебышевских многочленов позволяет поставить задачу и по-другому. Зададимся точностью ε (например, в нашем случае положим $\varepsilon \approx 1/5$). Тогда с помощью (1.44) можно “свернуть” ряд (1.42) так, что при той же точности он будет содержать гораздо меньше членов. Так как $|T_n(x)| \leq 1$, то отбрасывание трех последних членов в (1.44), содержащих $T_3(x), T_4(x)$ и $T_5(x)$, даст ошибку не более чем $\frac{7}{48} + \frac{1}{32} + \frac{1}{80} < \frac{1}{5,2}$. Отбрасывание члена $x^5/5$ в (1.41) также дает ошибку $< 1/5$. Тогда из (1.44) получим ряд

$$\ln(1+x) \approx -\frac{11}{32} + \frac{11}{8}x - \frac{3}{8}(2x^2 - 1) = \frac{1}{32} + \frac{11}{8}x - \frac{3}{4}x^2, \quad x \in [0, 1]. \quad (1.45)$$

С такой же точностью ($\varepsilon < 1/5$) ряд Тейлора (1.42) дает

$$\ln(1+x) \approx x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} \quad (1.46)$$

и степень многочлена в нем в два раза выше, чем у ряда (1.45), свернутого с помощью многочленов Чебышева. В табл. 1.1 для сравнения приведены точные значения функции $y = \ln(1+x)$, приближенные значения, полученные по формулам (1.45) и (1.46), и соответствующие значения погрешностей при значениях $x = 0,1; 0,5; 1,0$.

Таблица 1.1

x	Ряд Тейлора (1.46)	Свертка по Чебышеву (1.45)	Точное значение	Фактическая ошибка	
				для (1.46)	для (1.45)
0	0	0,03125	0	0	0,03125
0,1	0,09530633	0,1612500	0,0953102	0,00000387	0,06524
0,5	0,40104167	0,5312500	0,4054651	0,00442343	0,125785
1	0,5833333	0,656250	0,69314718	0,10981385	0,036897

Видно, что фактическая ошибка в чебышевском разложении распределена достаточно равномерно, в то время как у ряда Тейлора ошибка растет с ростом x . Очевидно, что при $x \approx 0$ ряд Тейлора имеет очень высокую точность, однако при $x \approx 1$ использование чебышевского разложения явно предпочтительнее. Так как при более высокой точности степень многочлена оказывается в 2 раза ниже, это приводит к значительному уменьшению объема общей вычислительной работы.

1.4. Среднеквадратичные приближения

Среднеквадратичное приближение – это способ аппроксимации, при котором критерий близости исходной функции $y = f(x)$ и аппроксимирующей ее функции $\varphi(x)$ определяется как сумма квадратов отклонений $f(x)$ от $\varphi(x)$ на заданной системе точек и ищется такая функция $\varphi(x)$; для которой эта сумма будет минимальна (метод наименьших квадратов). Такой подход особенно часто используется в том случае, когда исходная табличная функция $f(x)$ получена в результате экспериментов и требуется описать эту функцию аналитическим выражением, вид которого известен, но параметры, от которых зависит это выражение, подлежат определению. Другими словами, в качестве аппроксимирующей функции берется известная функция $\varphi(x; a_0, a_1, \dots, a_m)$, зависящая от значений параметров a_i . Пусть функция $y = f(x)$ задана таблично, т.е. задано дискретное множество аргументов $\{x_0, x_1, \dots, x_n\}$ и соответствующее множество значений функции $\{y_0, y_1, \dots, y_n\}$. Требуется подобрать такие значения параметров $a_0, a_1, \dots, a_m$, при которых сумма квадратов отклонений функций $f(x)$ и $\varphi(x)$ при всех значениях x_i ($i = 0, 1, \dots, n$)

$$S = \sum_{i=0}^n \eta_i^2 = \sum_{i=0}^n [\varphi(x_i; a_0, a_1, \dots, a_m) - y_i]^2 \quad (1.47)$$

будет минимальна. Поскольку значения аргументов x_i заданы, функция S будет зависеть только от параметров

$$S = S(a_0, a_1, \dots, a_m) \quad (1.48)$$

и задача сводится к нахождению минимума функции S . Из курса математического анализа известно, что для нахождения минимума

функции (1.48) надо найти частные производные этой функции по всем переменным a_i и приравнять их нулю

$$\frac{\partial S}{\partial a_0} = 0, \quad \frac{\partial S}{\partial a_1} = 0, \quad \dots, \quad \frac{\partial S}{\partial a_m} = 0. \quad (1.49)$$

Выражения (1.49) представляют собой систему $m + 1$ уравнения для определения $m + 1$ неизвестного параметра $a_0, a_1, \dots, a_m$. Если $S(a_0, a_1, \dots, a_m)$ зависит от аргументов линейным образом, то система (1.49) будет линейной, в противном случае – нелинейной, что значительно усложняет нахождение параметров a_i .

Для примера приведем использование метода наименьших квадратов для частного случая, когда в качестве аппроксимирующей функции выбран квадратный многочлен

$$\varphi(x) = a_0 + a_1x + a_2x^2, \quad (1.50)$$

в котором зависимость от коэффициентов a_0, a_1, a_2 является линейной. В этом случае выражение (1.47) примет вид

$$S = \sum_{i=0}^n (a_0 + a_1x_i + a_2x_i^2 - y_i)^2 = S(a_0, a_1, a_2). \quad (1.51)$$

Найдем частные производные (1.49) и приравняем их нулю:

$$\begin{aligned} \frac{\partial S}{\partial a_0} &= 2 \sum_{i=0}^n (a_0 + a_1x_i + a_2x_i^2 - y_i) = 0; \\ \frac{\partial S}{\partial a_1} &= 2 \sum_{i=0}^n (a_0 + a_1x_i + a_2x_i^2 - y_i)x_i = 0; \\ \frac{\partial S}{\partial a_2} &= 2 \sum_{i=0}^n (a_0 + a_1x_i + a_2x_i^2 - y_i)x_i^2 = 0. \end{aligned} \quad (1.52)$$

Раскрывая скобки и собирая коэффициенты при неизвестных, получим систему трех линейных уравнений для неизвестных a_0, a_1, a_2 :

$$\begin{cases} a_0(n+1) + a_1 \sum_{i=0}^n x_i + a_2 \sum_{i=0}^n x_i^2 = \sum_{i=0}^n y_i \\ a_0 \sum_{i=0}^n x_i + a_1 \sum_{i=0}^n x_i^2 + a_2 \sum_{i=0}^n x_i^3 = \sum_{i=0}^n x_i y_i \\ a_0 \sum_{i=0}^n x_i^2 + a_1 \sum_{i=0}^n x_i^3 + a_2 \sum_{i=0}^n x_i^4 = \sum_{i=0}^n x_i^2 y_i \end{cases} \quad (1.53)$$

Матрица полученной линейной системы симметрична, что удобно при проведении вычислений. Рассмотрим пример использования метода наименьших квадратов ([1]).

Пример 3. Функция $y = f(x)$ задана в табличном виде:

x	0,75	1,5	2,25	3,0	3,75
y	2,5	1,2	1,12	2,25	4,28

Найти квадратный трехчлен вида (1.50), который даст наилучшее среднеквадратичное приближение для функции $f(x)$.

Решение. В нашем случае $m = 2$, $n = 4$. Вычислим члены, являющиеся элементами матрицы системы (1.53) и ее правыми частями:

$$\sum_{i=0}^4 x_i = 0,75 + 1,50 + 2,25 + 3,00 + 3,75 = 11,25;$$

$$\sum_{i=0}^4 x_i^2 = 0,75^2 + 1,50^2 + 2,25^2 + 3,00^2 + 3,75^2 = 30,94;$$

$$\sum_{i=0}^4 x_i^3 = 0,75^3 + 1,50^3 + 2,25^3 + 3,00^3 + 3,75^3 = 94,92;$$

$$\sum_{i=0}^4 x_i^4 = 0,75^4 + 1,50^4 + 2,25^4 + 3,00^4 + 3,75^4 = 309,76;$$

$$\sum_{i=0}^4 y_i = 2,5 + 1,2 + 1,12 + 2,25 + 4,28 = 11,35;$$

$$\sum_{i=0}^4 x_i y_i = 29,00 \quad \sum_{i=0}^4 x_i^2 y_i = 90,21.$$

В нашем случае система (1.53) примет вид

$$\begin{cases} 5a_0 + 11,25a_1 + 30,94a_2 = 11,35 \\ 11,25a_0 + 30,94a_1 + 94,92a_2 = 29,00 \\ 30,94a_0 + 94,92a_1 + 309,76a_2 = 90,21 \end{cases} \quad (1.54)$$

Решив линейную систему (1.54), найдем: $a_0 = 5,54$; $a_1 = -4,73$; $a_2 = 1,19$. Следовательно, для заданной табличной функции получаем следующую аппроксимацию

$$y \approx 5,54 + 4,73x + 1,19x^2 = \varphi(x).$$

Можно оценить относительные погрешности полученной аппроксимации в заданных точках

$$\delta y_i = \frac{\varepsilon_i}{y_i} = \frac{\varphi(x_i) - y_i}{y_i}.$$

Результаты этой оценки приведены в табл. 1.2.

Таблица 1.2

x_i	y_i	$\varphi(x_i)$	ε_i	δy_i
0,75	2,50	2,66	0,16	0,064
1,50	1,20	1,12	-0,08	-0,067
2,25	1,12	0,92	-0,20	-0,179
3,00	2,25	2,06	-0,19	-0,084
3,75	4,28	4,54	0,26	0,061

2. ЧИСЛЕННОЕ ИНТЕГРИРОВАНИЕ

2.1. Постановка задачи

В этом разделе рассматриваются методы вычисления определенного интеграла

$$S = \int_a^b f(x) dx, \quad (2.1)$$

где $f(x)$ – подынтегральная функция, a и b – нижний и верхний пределы интегрирования.

Численное интегрирование – это вычисление определенного интеграла путем замены подынтегральной функции более простой аппроксимирующей функцией, последующего прямого интегрирования и получения расчетных формул (квадратурные формулы). Идея численного интегрирования заложена в самом определении определенного интеграла (2.1), которое вводится в курсе высшей математики. Напомним это определение. Пусть задана функция $y=f(x)$, определенная на отрезке $[a, b]$. Введем дискретное множество точек

$x_0, x_1, x_2, \dots, x_n$, причем $x_0 = a, x_n = b$. Эти точки разбивают отрезок $[a, b]$ на n элементарных отрезков $[x_{i-1}, x_i]$ ($i = 1, 2, \dots, n$). Обозначим через $h_i = x_i - x_{i-1}$ длину каждого элементарного отрезка. На каждом из этих отрезков выберем произвольную точку z_i ($x_{i-1} \leq z_i \leq x_i$) и найдем величину s_i

$$s_i = f(z_i) \cdot h_i. \quad (2.2)$$

Составим сумму всех полученных величин s_i для всех элементарных отрезков

$$S_n = s_1 + s_2 + \dots + s_n = \sum_{i=1}^n f(z_i) h_i. \quad (2.3)$$

Назовем интегральной суммой величину S_n . Обозначим через h длину наибольшего из отрезков: $h = \max(h_i)$. Тогда **определенным интегралом** от функции $f(x)$ на отрезке $[a, b]$ называется предел (если он существует) интегральной суммы при неограниченном увеличении

числа точек разбиения, при котором длина наибольшего из отрезков стремится к нулю

$$\int_a^b f(x)dx = \lim_{h \rightarrow 0} S_n = \lim_{h \rightarrow 0} \sum_{i=1}^n f(z_i)h_i. \quad (2.4)$$

Геометрический смысл рассмотренных понятий для случая $f(x) > 0$ показан на рис. 2.1.

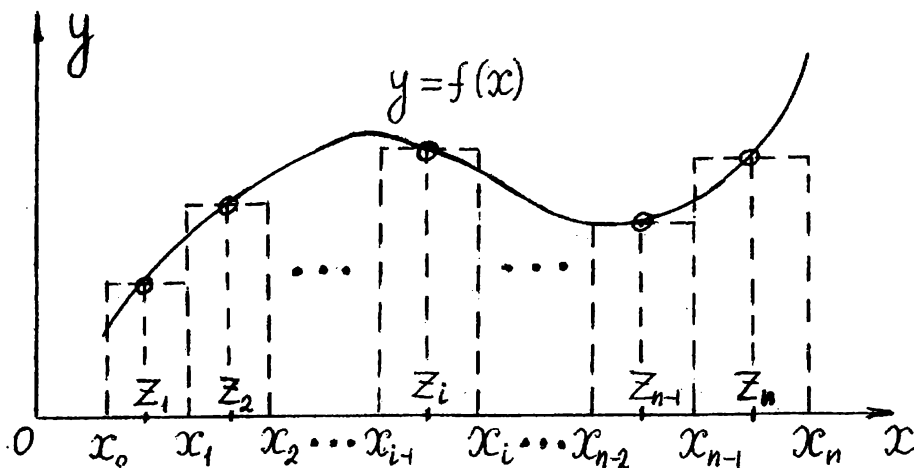


Рис. 2.1

Отмеченные кружками точки на графике функции $y = f(x)$ имеют абсциссы z_i и ординаты $f(z_i)$. Выражения вида (2.2) описывают площади элементарных прямоугольников (они отмечены штриховыми линиями), а интегральная сумма (2.3) – площадь ступенчатой фигуры, которая приближенно заменяет площадь криволинейной трапеции под кривой $y = f(x)$, ограниченной снизу осью x , а также прямыми $x = x_0 = a$ и $x = x_n = b$. Если предел (2.4) существует при $h \rightarrow 0$ и этот предел не зависит ни от способа разбиения на элементарные отрезки, ни от выбора точек z_i , то этот предел называется определенным интегралом от функции $f(x)$ на отрезке $[a, b]$. Отсюда следует, что геометрический смысл определенного интеграла – это площадь описанной выше криволинейной трапеции, а сам интеграл является вещественным числом. Функция $f(x)$, для которой существует предел

(2.4), называется интегрируемой. В частности, можно показать, что интегрируемой является любая непрерывная функция $f(x)$.

Различные формулы численного интегрирования, в сущности, основаны на соотношениях типа (2.3) и отличаются друг от друга следующим: выбором узловых точек x_i и точек z_i , в которых вычисляется функция $f(x)$; способом ускорения сходимости в процессе (2.4); величиной погрешностей и их оценкой, использующей дополнительную информацию о поведении функции $f(x)$.

Из курса высшей математики известно, что если для функций $f(x)$ можно аналитически найти первообразную функцию $F(x)$ такую, что $F'(x) = f(x)$, то определенный интеграл можно вычислить по формуле Ньютона – Лейбница

$$\int_a^b f(x)dx = F(b) - F(a). \quad (2.5)$$

Однако на практике воспользоваться формулой (2.5) удастся далеко не всегда, и приходится использовать численные методы для вычисления определенного интеграла. Такая необходимость возникает в двух случаях:

1. Подынтегральная функция $f(x)$ имеет такой вид, что первообразную $F(x)$ нельзя выразить через элементарные функции с помощью приемов и методов, используемых в высшей математике. Примером такого интеграла является часто используемый в теории вероятности интеграл вида

$$S = \int_a^b e^{-x^2} dx,$$

для которого нельзя в явном виде записать первообразную $F(x)$.

2. Подынтегральная функция $f(x)$ задана в табличном виде, когда задано множество значений ее аргументов $x_0, x_1, x_2, \dots, x_n$ и соответствующие им значения функции $y_0, y_1, y_2, \dots, y_n$.

Идея методов численного интегрирования основана на аппроксимации (приближении) подынтегральной функции $f(x)$ какой-либо близкой к ней аппроксимирующей функцией $\varphi(x)$, интеграл от которой может быть вычислен аналитически, т.е. может быть найдена в явном виде первообразная этой функции. Удобнее всего для этой цели использовать многочленные аппроксимации $\varphi(x) = P_m(x)$, которые были рассмотрены в первом разделе, так как интегралы от многочленов вычисляются элементарно. Например, **квадратурные формулы Ньютона-Котеса** получаются путем замены подынтегральной функции

интерполяционным многочленом Лагранжа с разбиением отрезка интегрирования на n равных частей. Часто встречающиеся на практике **квадратурные формулы прямоугольников, трапеций и Симпсона** – частные случаи формул Ньютона-Котеса, заменяющие подынтегральную функцию соответственно кусочно-постоянной, кусочно-линейной и кусочно-квадратичной функцией.

Рассмотрим вопросы, связанные с получением формул численного интегрирования. Интеграл (2.1) можно представить в виде суммы

интегралов по элементарным отрезкам $S = \sum_{i=1}^n S_i$, где

$$S_i = \int_{x_{i-1}}^{x_i} f(x) dx. \quad (2.6)$$

Замена интеграла (2.1) суммой интегралов вида (2.6) удобна тем, что позволяет использовать вместо глобальной интерполяции функции $f(x)$ на всем отрезке $[a, b]$ локальную интерполяцию на каждом из отрезков $[x_{i-1}, x_i]$ с помощью интерполяционных многочленов невысокой степени. Заменяв $f(x)$ многочленом $P_{mi}(x)$ на отрезке $[x_{i-1}, x_i]$, мы можем далее непосредственно вычислить интеграл

$$\overline{S}_i = \int_{x_{i-1}}^{x_i} P_{mi}(x), \quad (2.7)$$

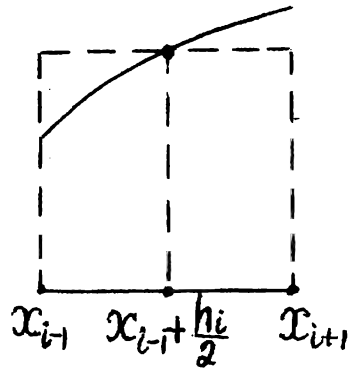
который с некоторой погрешностью, которую надо будет определить, будет равен интегралу (2.6)

$$S_i = \overline{S}_i + R_i. \quad (2.8)$$

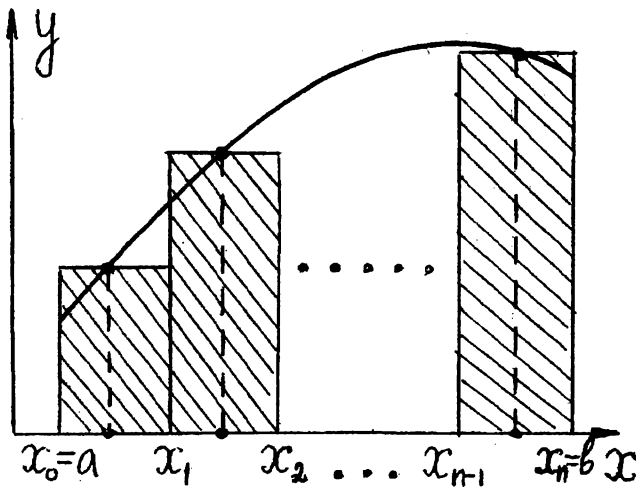
Здесь $P_{mi}(x)$ – интерполяционный многочлен степени m , который интерполирует подынтегральную функцию $f(x)$ на отрезке $[x_{i-1}, x_i]$, R_i – погрешность вычисления интеграла S_i . Задавая различные многочлены $P_{mi}(x)$ и вычисляя интеграл $\overline{S}_i$, мы получаем различные **квадратурные формулы**, которые приближенно аппроксимируют интеграл (2.6). Переходя затем к вычислению интеграла S на всем отрезке $[a, b]$, мы получаем **составную квадратурную формулу**. Приведем типичные квадратурные формулы, используемые обычно на практике.

2.2. Метод прямоугольников

Это простейший метод численного интегрирования, в котором подынтегральная функция $f(x)$ заменяется кусочно-постоянной функцией, т.е. функцией, постоянной на каждом отрезке $[x_{i-1}, x_i]$. Значение этой постоянной на каждом отрезке будет разным, равным $f(x_{i-1} + h_i/2)$ (рис. 2.2).



(a)



(б)

Рис. 2.2

На рис. 2.2 показан геометрический смысл такой замены на одном элементарном отрезке $[x_{i-1}, x_i]$ (рис. 2.2,а) и на всем отрезке $[a, b]$ (рис. 2.2,б). Видно, что искомая площадь заменяется суммой площадей соответствующих прямоугольников. На каждом элементарном отрезке

$$S_i = \int_{x_{i-1}}^{x_i} f(x)dx = f(x_{i-1} + h_i/2) \cdot h_i + R_i, \quad (2.9)$$

а на всем отрезке $[a, b]$

$$S = \int_a^b f(x)dx = \sum_{i=1}^n h_i \cdot f(x_{i-1} + h_i/2) + R. \quad (2.10)$$

Если пренебречь погрешностью R_i на элементарном отрезке и погрешностью R на всем отрезке, то (2.9) даст квадратурную формулу на элементарном отрезке, а (2.10) – составную квадратурную формулу для метода прямоугольников. В сущности, отличие (2.9) и (2.10) от формул (2.2) и (2.3) заключается только в специальном выборе точек z_i , а именно $z_i = x_{i-1} + 0,5h_i$, т.е. в качестве величины z_i выбирается середина отрезка $[x_{i-1}, x_i]$:

$$z_i = \frac{x_{i-1} + x_i}{2} = x_{i-1} + 0,5h_i. \quad (2.11)$$

Для применения квадратурных формул (2.9) и (2.10) необходимо оценить величины погрешностей R_i и R . Будем считать, что $f(x)$ непрерывна на $[x_{i-1}, x_i]$ вместе со своей первой и второй производными. Вводя для функции $f(x)$ ее первообразную $F(x)$, для которой $F'(x) = f(x)$, запишем левую часть (2.9) в виде

$$\int_{x_{i-1}}^{x_i} f(x)dx = F(x_i) - F(x_{i-1}). \quad (2.12)$$

Разлагая $F(x)$ в ряд Тейлора в окрестности базовой точки $x = z_i$ с остаточным членом в форме Лагранжа, получим значения функции $F(x_i)$ и $F(x_{i-1})$:

$$F(x_i) = F(z_i) + \frac{h}{2}F'(z_i) + \left(\frac{h}{2}\right)^2 \cdot \frac{1}{2!}F''(z_i) + \left(\frac{h}{2}\right)^3 \cdot \frac{1}{3!}F'''(\zeta_1), \quad (2.13)$$

$$F(x_{i-1}) = F(z_i) - \frac{h}{2}F'(z_i) + \left(\frac{h}{2}\right)^2 \cdot \frac{1}{2!}F''(z_i) - \left(\frac{h}{2}\right)^3 \cdot \frac{1}{3!}F'''(\zeta_2), \quad (2.14)$$

где $\zeta_1 \in [z_i, x_i]$, а $\zeta_2 \in [x_{i-1}, z_i]$. Подставляя в правую часть (2.12) выражения (2.13) и (2.14) и учитывая, что $F'(x) = f(x)$, получим

$$\int_{x_{i-1}}^{x_i} f(x)dx = hf(z_i) + \frac{h^3}{24}f''(\zeta), \quad (2.15)$$

где $f''(\zeta) = 0,5[f''(\zeta_1) + f''(\zeta_2)]$ [такое значение ζ существует в силу непрерывности функции $f(x)$]. Последний член в (2.15) дает выражение для погрешности квадратурной формулы метода прямоугольников

$$R_i = \frac{h^3}{24}f''(\zeta). \quad (2.16)$$

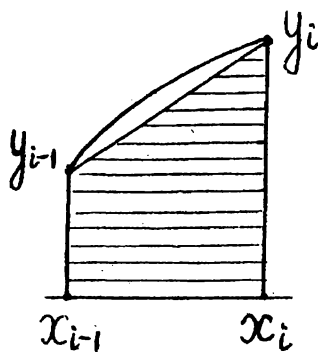
Для составной квадратурной формулы (2.10) вычисления интеграла на всем отрезке $[a, b]$ погрешность имеет вид

$$R = \frac{h^2(b-a)}{24}f''(\zeta) \quad (\zeta \in [a, b]). \quad (2.16a)$$

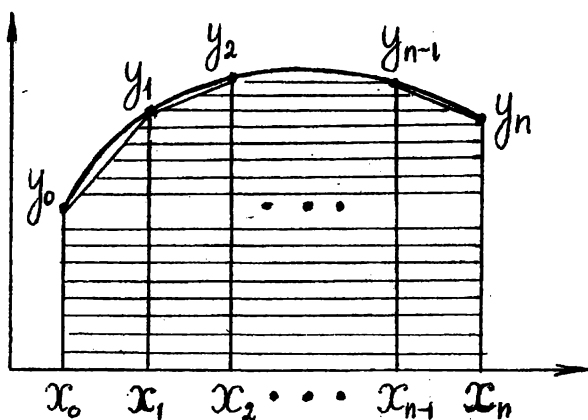
Видно, что для функции с ограниченной второй производной погрешности R_i и R убывают с уменьшением шага разбиения h . При малых h R_i имеют третий порядок малости, а R – второй. Это объясняется тем, что при суммировании всех погрешностей R_i на всем отрезке $[a, b]$ величина (2.16) умножается на количество элементарных отрезков n , которое при $h_i = \text{const} = h$ связано с величиной h соотношением $n = (b - a)/h$.

2.3. Метод трапеций

Метод трапеций использует для описания функции $f(x)$ линейную интерполяцию, заменяя ее кусочно-линейной функцией на отрезке $[a, b]$, т.е. ломаной, соединяющей точки x_i , $y_i = f(x_i)$ (рис. 2.3)



(a)



(б)

Рис. 2.3

В этом случае искомая площадь заменяется суммой трапеций, одна из которых приведена на рис. 2.3,а. Площадь этой трапеции легко вычисляется и дает квадратурную формулу

$$S_i = \int_{x_{i-1}}^{x_i} f(x)dx = \frac{1}{2}(y_{i-1} + y_i)h_i + R_i. \quad (2.17)$$

Методом, аналогичным приведенному выше, можно показать, что локальная погрешность метода трапеций имеет вид

$$R_i = -\frac{1}{12}h_i^3 f''(\zeta_i) \quad (\zeta \in [x_{i-1}, x_i]), \quad (2.18)$$

а погрешность на всем интервале $[a, b]$ равна

$$R = -\frac{h^2(b-a)}{12} f''(\zeta) \quad (\zeta \in [a, b]). \quad (2.19)$$

Сравнение погрешностей R_i и R для метода трапеций [формулы (2.18) и (2.19)] и метода прямоугольников [формулы (2.16) и (2.16а)] показывает, что эти погрешности имеют разные знаки, и погрешность метода прямоугольников в 2 раза меньше (с точностью до величины второй производной $f''(\zeta)$ в разных, вообще говоря, точках ζ для разных методов). Разные знаки у формул (2.16а) и (2.19) указывают на то, что методы прямоугольников и трапеций дают двусторонние приближения к решению. Это означает, что если один из методов дает решение с недостатком, то другой метод – с избытком: точное решение находится в вилке между значениями, полученными методами прямоугольников и трапеций, что дает надежную оценку для величины определенного интеграла.

Запишем теперь составную квадратурную формулу для метода трапеций

$$S = \int_a^b f(x)dx = \sum_{i=1}^n \frac{1}{2}(y_{i-1} + y_i) \cdot h_i + R. \quad (2.20)$$

В случае постоянного шага $h_i = x_i - x_{i-1} = \text{const} = h$ формулу метода трапеций (2.20) можно без учета погрешностей записать в виде

$$\begin{aligned} S &= 0,5h [(y_0 + y_1) + (y_1 + y_2) + (y_2 + y_3) + \dots + (y_{n-1} + y_n)] = \\ &= h [0,5(y_0 + y_n) + y_1 + y_2 + \dots + y_{n-1}]. \end{aligned} \quad (2.21)$$

Рассмотрим применение метода прямоугольников и трапеций на примере.

Пример 4. Вычислить интеграл

$$S = \int_0^1 \frac{dx}{1+x^2}.$$

Решение. Этот интеграл является табличным и легко вычисляется по формуле Ньютона – Лейбница

$$\bar{S} = (\arctg(x)) \Big|_0^1 = \pi/4 = 0,785398.$$

Применим для вычисления интеграла метод прямоугольников [формула (2.10)] и метод трапеций [формула (2.21)]. Разобьем отрезок интегрирования $[0, 1]$ на пять равных частей: $n = 5$, $h = 0,2$. Вычислим значения подынтегральной функции $y_i = 1/(1+x_i^2)$ в точках разбиения $x_i = x_{i-1} + 0,2$ (для метода трапеций) и в серединах отрезков разбиения $z_i = x_{i-1} + 0,5h$ (для метода прямоугольников). Результаты вычислений представлены в табл. 2.1.

Таблица 2.1

x_i	$y_i = y(x_i)$	$z_i = x_{i-1} + 0,5h$	$y(z_i)$	
0,0	1,000000	0,1	0,990099	Метод прямоугольников $S_{\Pi} = 0,786237$
0,2	0,961538	0,3	0,917431	
0,4	0,862069	0,5	0,800000	
0,6	0,735294	0,7	0,671141	Метод трапеций $S_{\tau} = 0,783731$
0,8	0,609756	0,9	0,552486	
1,0	0,500000			Точное решение $\bar{S} = 0,785398$

Методом прямоугольников получим [по формуле (2.10)]

$$\begin{aligned} S_{\Pi} &= 0,2(0,990099 + 0,917431 + 0,8 + 0,671141 + 0,552486) = \\ &= 0,2 \cdot 3,931157 = 0,786231. \end{aligned}$$

Методом трапеций [по формуле (2.21)]

$$S_T = 0,2 [(1+0,5) \cdot 0,5 + 0,961538 + 0,862069 + 0,735294 + 0,609756] = \\ = 0,2 \cdot 3,918657 = 0,783731.$$

Результаты показывают, что метод прямоугольников дает завышенное значение интеграла, а метод трапеций – заниженное, как это и следует из сравнения погрешностей этих методов (2.16а) и (2.19). Фактические значения погрешностей составляют: для метода прямоугольников $\bar{S} - S_{\Pi} = -0,000833$, а для метода трапеций $\bar{S} - S_T = 0,001667$. Погрешности для метода прямоугольников оказываются почти в 2 раза меньше, что соответствует теоретическим оценкам (2.16а) и (2.19). На первый взгляд, более высокая точность метода прямоугольников является неожиданной, так как он использует более грубую кусочно-постоянную аппроксимацию функции $f(x)$, чем метод трапеций, использующий кусочно-линейную интерполяцию. Дело здесь заключается в выборе точки $z_i = 0,5(x_{i-1} + x_i)$, в которой вычисляется функция $f(x)$. Если в качестве z_i выбирать левый конец интервала $z_i = x_{i-1}$ или его правый конец $z_i = x_i$, то мы получим другие варианты метода прямоугольников, которые будут давать гораздо большую погрешность. Соответствующие формулы имеют вид:

$$S = \sum_{i=1}^n y_{i-1} h_i; \quad (2.22)$$

$$S = \sum_{i=1}^n y_i h_i. \quad (2.23)$$

Геометрический смысл применения этих формул изображен на рис. 2.4, где пунктиром показаны площади прямоугольников, которые приближают площадь криволинейной трапеции при различных вариантах метода прямоугольников. Единственное достоинство методов (2.22) и (2.23) заключается в том, как это видно на рис. 2.4, что функция $f(x)$ вычисляется только в выбранных точках разбиения x_i , в то время как более точный метод (2.10) требует вычисления функции в середине отрезка $[x_{i-1}, x_i]$. Обычно вычисление по формуле (2.22) называют

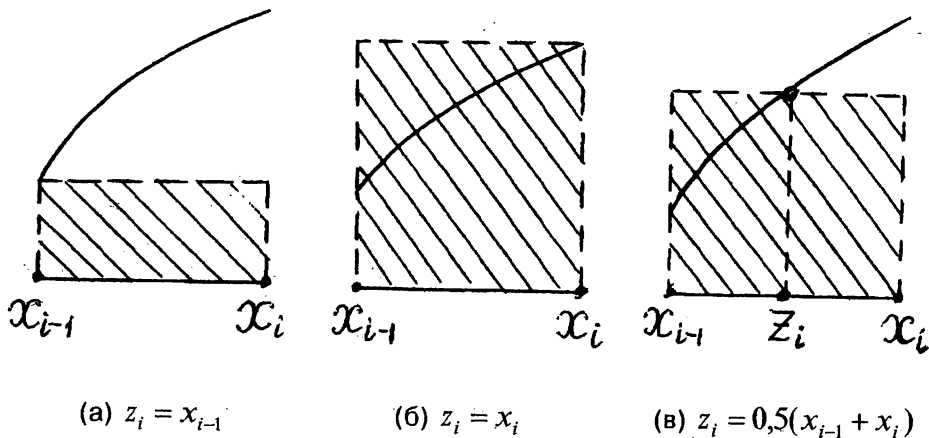


Рис. 2.4

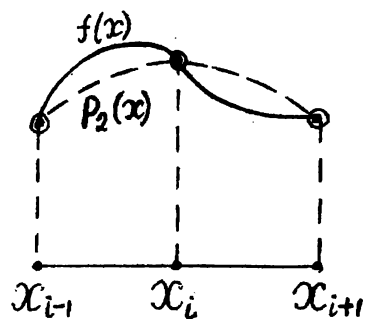
методом прямоугольников по левой границе, а по (2.23) – методом прямоугольников по правой границе.

2.4. Метод Симпсона

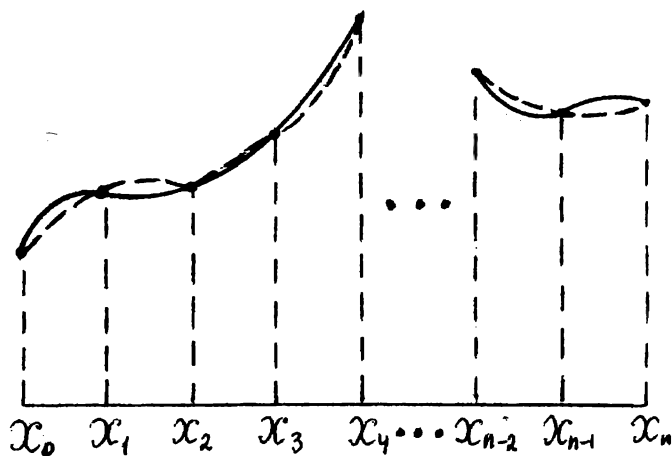
Во многих случаях требуется вычислять интегралы для таблично заданных функций, когда точки x_i разбиения отрезка $[a, b]$ на элементарные отрезки заданы и повысить точность численного интегрирования путем уменьшения шага h_i не удастся. В этом случае желательно использовать квадратурные формулы, которые при том же разбиении отрезка дают более высокую точность, чем методы прямоугольников и трапеций. Одним из таких методов является метод Симпсона. Повышение точности в нем достигается за счет повышения точности интерполяции – функция $f(x)$ заменяется интерполяционным многочленом Лагранжа второй степени (квадратичная интерполяция).

Разобьем отрезок интегрирования $[a, b]$ на четное число равных частей, полагая $x_{i+1} - x_i = x_i - x_{i-1} = h$. Рассмотрим два соседних интервала $[x_{i-1}, x_i]$ и $[x_i, x_{i+1}]$ (рис. 2.5,а).

Проведем через три точки $(x_{i-1}, y_{i-1}), (x_i, y_i), (x_{i+1}, y_{i+1})$ интерполяционный многочлен Лагранжа второй степени $L_2(x)$ [формула (1.15)]. Вычислим определенный интеграл на отрезке $[x_{i-1}, x_{i+1}]$, заменив



(a)



(б)

Рис. 2.5

подынтегральную функцию $f(x)$ многочленом $L_2(x)$. После интегрирования получим

$$S_i = \int_{x_{i-1}}^{x_{i+1}} f(x) dx \approx \int_{x_{i-1}}^{x_{i+1}} L_2(x) dx = \frac{h}{3} (y_{i-1} + 4y_i + y_{i+1}). \quad (2.24)$$

Формула (2.24) является квадратурной формулой метода Симпсона на отрезке $[x_{i-1}, x_{i+1}]$. Рассматривая теперь весь отрезок интегрирования $[a, b]$ и записывая (2.24) для попарно расположенных отрезков $[x_0, x_2], [x_2, x_4], \dots, [x_{n-2}, x_n]$, можно получить составную квадратурную формулу метода Симпсона

$$S = \int_a^b f(x) dx \approx$$

$$\approx \frac{h}{3} [(y_0 + 4y_1 + y_2) + (y_2 + 4y_3 + y_4) + \dots + (y_{n-2} + 4y_{n-1} + y_n)] = \quad (2.25)$$

$$= \frac{h}{3} [y_0 + 4y_1 + 2y_2 + 4y_3 + 2y_4 + \dots + 2y_{n-2} + 4y_{n-1} + y_n].$$

В формуле (2.25) члены в квадратных скобках сгруппированы следующим образом: крайние члены $[y_0 = f(a) \text{ и } y_n = f(b)]$ входят в сумму с коэффициентом, равным единице; члены с четными номерами $(y_2, y_4, y_6, \dots, y_{n-2})$ умножаются на два, а члены с нечетными номерами $(y_1, y_3, y_5, \dots, y_{n-1})$ умножаются на четыре. Исследование погрешности метода Симпсона показывает, что локальная погрешность на отрезке $[x_{i-1}, x_{i+1}]$ имеет вид

$$R_i = -\frac{h^5}{90} f^{(4)}(\zeta_i) \quad (\zeta_i \in [x_{i-1}, x_{i+1}]), \quad (2.26)$$

а погрешность формулы (2.24) на всем отрезке $[a, b]$

$$R = -\frac{h^4(b-a)}{180} f^{(4)}(\zeta) \quad (\zeta \in [a, b]), \quad (2.27)$$

где $f^{(4)}(\zeta)$ – четвертая производная функции $f(x)$ в некоторой точке $x = \zeta$. Видно, что для достаточно гладких функций, когда $f^{(4)}(x)$ ограничена, метод Симпсона является гораздо более точным, чем методы прямоугольников и трапеций. Вычислим с его помощью определенный интеграл, рассмотренный в предыдущем подразделе.

Пример 5. Вычислить методом Симпсона интеграл

$$S = \int_0^1 \frac{dx}{1+x^2}.$$

Решение. Разобьем отрезок $[0, 1]$ на четыре элементарных отрезка: $n = 4$, $h = 0,25$. Результаты вычисления подынтегральной функции $y_i = 1/(1+x_i^2)$ приведены в табл. 2.2.

Таблица 2.2

x_i	y_i	Точное решение
0,0	1,000000	$\bar{S} = 0,785398$
0,25	0,941176	Метод Симпсона $S_c = 0,785392$
0,50	0,800000	
0,75	0,640000	Погрешность $\bar{S} - S_c = 0,000006$
1,0	0,500000	

По формуле Симпсона (2.25) получим

$$S_c = \frac{0,25}{3} (1 + 4 \cdot 0,941176 + 2 \cdot 0,8 + 4 \cdot 0,64 + 0,5) =$$

$$\frac{0,25}{3} \cdot 9,424704 = 0,785392.$$

Сравнение с результатами, приведенными в табл. 2.1, показывает, что хотя шаг h в данном случае выбран даже несколько больше ($h = 0,25$ вместо $h = 0,2$ в примере 4), погрешность метода Симпсона в 278 раз меньше, чем погрешность метода трапеций.

2.5. Метод Гаусса

Ранее уже отмечалось, что формулы прямоугольников, трапеций и Симпсона являются частным случаем квадратурных формул Ньютона-Котеса. Эти формулы получаются путем разбиения отрезков

интегрирования на n равных элементарных отрезков путем введения равноотстоящих узлов $x_i = ih$ ($i = 0, 1, 2, \dots, n$), где h – расстояние между узлами ($h = 1/n$). Далее строится интерполяционный многочлен Лагранжа $L_n(x)$, который совпадает с подынтегральной функцией в узловых точках. Интегрируя аналитически этот многочлен, мы получаем соответствующие квадратурные формулы. Формулы Ньютона-Котеса имеют вид

$$\int_0^{nh} f(x) dx = w_0 f(0) + w_1 f(h) + w_2 f(2h) + \dots + w_n f(nh). \quad (2.28)$$

Здесь $f(0), f(h), \dots, f(nh)$ – значения подынтегральной функции в узлах интегрирования, а $w_0, w_1, \dots, w_n$ – весовые коэффициенты, определяемые при интегрировании интерполяционного многочлена $L_n(x)$. Например, при $n = 1$ получается квадратурная формула трапеции

$$\int_0^h f(x) dx = \frac{h}{2} [f(0) + f(h)] - \frac{1}{12} h^3 f''(\zeta), \quad (2.29)$$

при $n = 2$ – формула Симпсона

$$\int_0^{2h} f(x) dx = \frac{h}{3} [f(0) + 4f(h) + f(2h)] - \frac{1}{90} h^5 f^{(4)}(\zeta). \quad (2.30)$$

Однако, как мы уже видели в разделе 1, интерполяция на равноотстоящих узлах имеет существенный недостаток, который проявляется при увеличении количества узлов (при возрастании n). Аналогичный недостаток существует и у формул Ньютона-Котеса (2.28). В частности, он проявляется в том, что при увеличении количества узлов, в которых вычисляется функция $f(x)$, не увеличивается точность вычисления интеграла даже для очень гладких функций $f(x)$. Этот недостаток называют явлением насыщения численного метода. Таким недостатком не обладает метод Гаусса, который будет здесь рассмотрен.

Главная идея метода Гаусса заключается в том, что при заданном количестве n разбиений интервала $[a, b]$ можно попытаться расположить узлы разбиений x_i ($i = 0, 1, \dots, n$) не на одинаковом расстоянии друг от

друга, а подобрать такое их расположение, при котором точность вычисления интеграла будет наибольшей. Если записать квадратурную формулу в общем виде

$$\int_a^b f(x)dx \approx \sum_{i=0}^n f(x_i)w_i, \quad (2.31)$$

то задача заключается в том, чтобы при заданном значении n подобрать такое расположение узлов x_i и выбрать такие соответствующие весовые коэффициенты w_i , чтобы погрешность была наименьшей. Такая задача была поставлена и решена известным немецким математиком Гауссом. Полученная Гауссом квадратурная формула не обладает свойством насыщения и имеет очень высокую точность. Не рассматривая здесь способ получения формулы метода Гаусса, приведем лишь окончательный результат, который позволяет использовать метод Гаусса в практических вычислениях.

При вычислении интеграла $\int_a^b f(x)dx$ прежде всего необходимо с помощью замены переменной преобразовать интеграл так, чтобы отрезком интегрирования был отрезок $[-1, 1]$, симметричный относительно начала координат. Этого всегда можно добиться, переходя к новой переменной ξ с помощью преобразования

$$x = \frac{b+a}{2} + \frac{b-a}{2}\xi. \quad (2.32)$$

Действительно, при $x = a$ из (2.32) получаем

$$a = \frac{b}{2} + \frac{a}{2} + \left(\frac{b}{2} - \frac{a}{2}\right)\xi \Rightarrow \frac{a}{2} - \frac{b}{2} = \left(\frac{b}{2} - \frac{a}{2}\right)\xi \Rightarrow \xi = -1.$$

При $x = b$

$$b = \frac{b}{2} + \frac{a}{2} + \left(\frac{b}{2} - \frac{a}{2}\right)\xi \Rightarrow \frac{b}{2} - \frac{a}{2} = \left(\frac{b}{2} - \frac{a}{2}\right)\xi \Rightarrow \xi = 1.$$

Таким образом, после преобразования (2.32) исходный интеграл примет вид

$$S = \int_a^b f(x)dx = \int_{-1}^1 \varphi(\xi)d\xi. \quad (2.33)$$

Значение S определенного интеграла при этом, конечно, не изменится. К интегралу, стоящему в правой части, можно применять метод Гаусса. Гаусс показал, что наилучшая по точности квадратура для (2.33) имеет вид

$$\int_{-1}^1 \varphi(\xi)d\xi = \sum_{i=1}^n \varphi(\xi_i)w_i + R_g. \quad (2.34)$$

Здесь ξ_i ($i = 0, 1, \dots, n$) – корни многочлена Лежандра $l_n(x)$ степени n . Эти корни являются простыми (нет кратных, т.е. совпадающих корней: $\xi_i \neq \xi_j$ при $i \neq j$), все корни являются вещественными числами и располагаются на отрезке $[-1, 1]$ симметрично относительно начала координат $\xi = 0$. Коэффициенты w_i определяются путем вычисления специальных определенных интегралов. Количество этих коэффициентов равно степени n многочлена Лежандра $l_n(x)$. R_g – погрешность квадратурной формулы Гаусса.

Относительно вида формулы Гаусса (2.34) следует сделать несколько пояснений.

1. В ней используются значения корней многочленов Лежандра $l_n(x)$. Эти многочлены, как и описанные в первом разделе многочлены Чебышева, относятся к семейству специальных многочленов, которые широко используются в математике и в численных методах. В первой главе отмечалось, что из всех многочленов степени n со старшим коэффициентом (при x^n), равным единице, наименее уклоняющимся от нуля является многочлен Чебышева, если рассматривать равномерное приближение функции $f(x)$ на всем отрезке $[-1, 1]$. Это означает, что многочлен Чебышева дает наилучшее равномерное приближение для $f(x)$ по сравнению со всеми другими многочленами той же степени. Оказывается, что многочлен Лежандра дает наилучшее приближение среди всех многочленов, если рассматривать среднеквадратичное приближение (приближение в “среднем”). Приведем несколько первых многочленов $l_n(x)$:

$$l_0(x)=1, \quad l_1(x)=x, \quad l_2(x)=\frac{1}{3}(3x^2-1), \quad l_3(x)=\frac{1}{5}(5x^3-3x).$$

Корни многочленов Лежандра ξ_i ($i = 0, 1, \dots, n$) для любого заданного значения n можно найти во многих книгах по численным методам. Например, в работе [7] даются их значения до $n = 9$.

2. Корни ξ_i многочленов Лежандра $L_n(x)$ и коэффициенты w_i зависят только от степени многочлена n и не зависят от вида подынтегральной функции $\varphi(\xi)$ в (2.34). Это позволяет составить таблицы величин ξ_i и w_i в зависимости от значения n и использовать их при численном интегрировании методом Гаусса. Приведем несколько первых значений этих величин, округленных до шести цифр после запятой (табл. 2.3).

Отметим, что в отличие от методов трапеций и Симпсона, где в квадратурной формуле используются значения функции $f(x)$ на границах интервала $[a, b]$ ($f(a)$ и $f(b)$), в квадратуре Гаусса граничные точки $\xi = -1$ и $\xi = 1$ не используются. Середина отрезка $\xi = 0$ используется только при нечетных значениях $n = 3, 5, 7, 9, \dots$. Все весовые коэффициенты w_i

Таблица 2.3

n	ξ_i	w_i
2	$\pm 0,577350$	1
3	0	0,888889
	$\pm 0,774597$	0,555556
4	$\pm 0,339981$	0,652145
	$\pm 0,861136$	0,347855
5	0	0,568889
	$\pm 0,538469$	0,478629
	$\pm 0,906180$	0,236927

положительны, а все корни ξ_i располагаются симметрично относительно точки $\xi = 0$. Весовые коэффициенты w_i для симметрично расположенных корней одинаковы и указываются в таблице один раз для каждой пары симметрично расположенных корней. С помощью таблицы 2.3 легко записать квадратурную формулу Гаусса для заданного количества узлов интегрирования. Например, для $n = 2$ $\xi_1 = 0,577350$; $\xi_2 = -0,577350$; $w_1 = w_2 = 1$ и квадратурная формула (2.34) примет вид (если пренебречь погрешностью R_g)

$$\int_{-1}^1 f(\xi) d\xi = f(-0,577350) + f(0,577350). \quad (2.35)$$

Для $n = 3$ получим

$$\int_{-1}^1 f(\xi) d\xi = [f(-0,774597) + f(0,774597)] \cdot 0,555556 + f(0) \cdot 0,888889. \quad (2.36)$$

Для использования формул (2.35), (2.36) и аналогичных формул при других n достаточно вычислить значения подынтегральной функции $f(\xi)$ в точках ξ_i , умножить их на свои весовые коэффициенты и сложить все полученные величины.

3. Точность квадратурной формулы Гаусса является очень высокой. Точная теоретическая оценка является достаточно сложной и использует значение производной порядка $2n - f^{(2n)}(\theta)$, где $\theta \in [-1, 1]$. Здесь лишь достаточно заметить, что, взяв n точек разбиения на отрезке интегрирования, мы получим, что квадратурная формула Гаусса будет точной для любого многочлена степени $2n$. Для сравнения укажем, что квадратурная формула трапеций является точной для $f(x)$, имеющей вид многочлена первой степени. Действительно, в выражение для ее погрешности входит величина $f''(\zeta)$ [формула (2.18)], которая равна нулю для многочлена первой степени. Квадратурная формула метода Симпсона является точной для функции $f(x)$, имеющей вид кубического многочлена, так как ее погрешность [формула (2.26)] зависит от величины $f^{(4)}(\zeta)$, которая равна нулю для кубического многочлена. Однако здесь надо учесть, что метод Гаусса имеет высокую точность только для очень гладких функций, имеющих достаточное количество ограниченных производных.

Рассмотрим применение метода Гаусса на том же примере, который в предыдущих подразделах иллюстрировал методы прямоугольников, трапеций и Симпсона.

Пример 6. Вычислить методом Гаусса определенный интеграл

$$S = \int_0^1 f(x) dx = \int_0^1 \frac{dx}{1+x^2}. \quad (2.37)$$

Решение. Можно сделать замену переменных (2.32) и получить в переменной ξ пределы интегрирования от -1 до $+1$, как это необходимо

для метода Гаусса. Но в данном случае можно поступить проще. Рассмотрим интеграл

$$P = \int_{-1}^1 \frac{dx}{1+x^2}, \quad (2.38)$$

величина которого будет вдвое больше величины S , так как

$f(x) = \frac{1}{1+x^2}$ – функция четная, в силу чего $P = 2S$. Вычислим интеграл

методом Гаусса, взяв на отрезке $[-1, 1]$ пять узловых точек. Из табл. 2.3

возьмем для случая $n = 5$ значения корней многочлена Лежандра $L_5(x)$

$\xi_1 = -0,906180$; $\xi_2 = -0,538469$; $\xi_3 = 0$; $\xi_4 = 0,538469$; $\xi_5 = 0,906180$

и соответствующие им коэффициенты w_i (для симметрично располо-

женных относительно $\xi = 0$ корней коэффициенты w_i равны). Из (2.34)

получим для этого случая следующую квадратурную формулу

$$P = \int_{-1}^1 f(x)dx \approx [f(-0,906180) + f(0,906180)] \cdot 0,236927 + \\ + [f(-0,538469) + f(0,538469)] \cdot 0,478629 + f(0) \cdot 0,568889. \quad (2.39)$$

Вычисляя функцию $f(x) = \frac{1}{1+x^2}$ в указанных точках, получим $P =$

1,571171. Для сравнения проведем расчет методом Симпсона для такого же количества узлов, расположив их равномерно на отрезке $[-1, 1]$: $x_1 = -1$; $x_2 = -0,5$; $x_3 = 0$; $x_4 = 0,5$; $x_5 = 1$. Квадратурная формула метода Симпсона в этом случае имеет вид

$$P = \int_{-1}^1 f(x)dx \approx \frac{h}{3} [0,5(f(-1) + f(1)) + 4(f(-0,5) + f(0,5)) + 2f(0)] = \\ = \frac{0,5}{3} [0,5(0,5 + 0,5) + 4(0,8 + 0,8) + 2 \cdot 1] = 1,566667.$$

Для величины S получим $S = 0,783333$, откуда погрешность метода Симпсона будет равна $R_s = \bar{S} - S = 0,00207$. Эти результаты показывают, что при одном и том же количестве узлов $n = 5$ погрешность квадратурной формулы Гаусса оказалась в нашем случае почти в 10 раз меньше, чем у формулы Симпсона, хотя и метод Симпсона имеет очень высокую точность.

2.6. Сравнительная оценка методов и способы уточнения решения

В рассмотренных выше методах численного интегрирования мы видели, что погрешности квадратурных формул зависят от шага интегрирования h и от значения производных k -го порядка в некоторой внутренней точке интервала [формулы (2.16а), (2.19), (2.27)]. Для методов прямоугольников и трапеций $k = 2$, для метода Симпсона $k = 4$, а для метода Гаусса $k = 2n$, где n – количество интервалов разбиения отрезка. Отсюда следует, что возможность применения того или иного метода зависит от гладкости функции, т.е. от существования и величины соответствующих производных. Для примера сравним погрешности метода трапеций (2.19) и Симпсона (2.27):

$$R = \frac{h^2(b-a)}{24} f''(\zeta_1); \quad R = -\frac{h^4(b-a)}{180} f^{(4)}(\zeta_2).$$

Ясно, что если значения $f''(\zeta_1)$ и $f^{(4)}(\zeta_2)$ ^{*)} имеют один порядок величины, то метод Симпсона явно предпочтительнее. Однако возможны случаи, когда $f^{(4)}(\zeta) \gg f''(\zeta)$ и погрешность метода Симпсона

может оказаться больше, если $\left| \frac{h^2 f^{(4)}(\zeta)}{180} \right| > \left| \frac{f''(\zeta)}{12} \right|$. Очевидно, что в

каждом конкретном случае для заданной подынтегральной функции необходимо проводить оценку соответствующих производных и решать, какой из методов предпочтительнее.

К сожалению, во многих практических задачах оценить величины

$M_2 = \max_{a \leq x \leq b} |f''(x)|$, $M_4 = \max_{a \leq x \leq b} |f^{(4)}(x)|$ либо очень трудно, либо вообще невозможно. В этом случае можно использовать практические способы

^{*)} ζ_1 и ζ_2 в общем случае различны, но для простоты будем обозначать их без индексов.

оценки погрешности численного интегрирования путем вычисления двух значений одного и того же интеграла с помощью двух семейств узлов. Одним из таких распространенных способов является **метод Рунге** – способ повышения точности численного интегрирования путем проведения расчетов с шагами h и $h/2$ и соответствующей корректировки результатов. Метод основан на использовании формул для погрешности (2.16а), (2.19) и (2.27). Формулы для вычисления интеграла можно записать в общем виде

$$S = S_h + ch^k + O(h^{k+1}), \quad h \rightarrow 0. \quad (2.40)$$

Здесь S – точное значение интеграла, S_h – приближенное значение интеграла, полученное каким-либо методом при равномерном разбиении отрезка $[a, b]$ с шагом h ; ch^k – погрешность интегрирования; k – порядок погрешности; c – коэффициент при h^k , равный соответствующему значению производной; $O(h^{k+1})$ – член более высокого порядка малости по сравнению с главным членом погрешности ch^k . Сравнение общей записи (2.40) с аналогичной записью для конкретных квадратурных формул показывает, что, например, для метода

трапеций $k = 2$, $c = -\frac{b-a}{12} f''(\zeta)$, а для метода Симпсона $k = 4$,

$$c = -\frac{b-a}{180} f^{(4)}(\zeta).$$

Предположим, что величина c не изменяется при вычислении интеграла с различными шагами h и что $c \neq 0$. Тогда при вычислении с шагом h имеет место соотношение (2.40), а при вычислении с вдвое меньшим шагом $h/2$ (2.40) примет вид

$$S = S_{h/2} + c \left(\frac{h}{2} \right)^k + O(h^{k+1}). \quad (2.41)$$

Поскольку для каждого конкретного метода величина k известна, мы можем, посчитав S_h и $S_{h/2}$, вычислить неизвестную константу c . Действительно, вычитая почленно (2.41) из (2.40), получим

$$0 = S_h - S_{h/2} + c \left(h^k - \frac{h^k}{2^k} \right) \Rightarrow c = \frac{(S_{h/2} - S_h) \cdot 2^k}{h^k (2^k - 1)}. \quad (2.42)$$

Подставляя теперь полученное значение c в соотношение (2.41), получим уточненное по методу Рунге решение $\bar{S}$

$$\bar{S} = S_{h/2} + \frac{S_{h/2} - S_h}{2^k - 1} + O(h^{k+1}). \quad (2.43)$$

Смысл полученного соотношения (2.43) заключается в следующем: если для погрешности квадратурной формулы порядок k известен, то, вычислив приближенное значение интеграла S_h с шагом h и $S_{h/2}$ с вдвое меньшим шагом $h/2$, мы можем приближенно вычислить главный член погрешности метода [второй член в правой части (2.43)] и, прибавив его к величине $S_{h/2}$, получить уточненное значение интеграла, более точное, чем величина $S_{h/2}$. Для метода трапеций ($k = 2$) получим, отбрасывая малый член $O(h^{k+1})$,

$$\bar{S} = S_{h/2} + \frac{S_{h/2} - S_h}{3}, \quad (2.44)$$

а для формулы Симпсона ($k = 4$)

$$\bar{S} = S_{h/2} + \frac{S_{h/2} - S_h}{15}. \quad (2.45)$$

Рассмотрим пример на использование правила Рунге.

Пример 7. Вычислить методом трапеций интеграл

$$S = \int_0^1 \frac{dx}{1+x^2},$$

посчитав его с шагом $h = 0,5$ и с шагом $h = 0,25$, и уточнить решение методом Рунге.

Решение. Значение подынтегральной функции в используемых узловых точках приведем в табл. 2.4.

Таблица 2.4

x_i	y_i	
0,00	1,000000	$S_{0,5} = 0, \underline{775}$ (решение с шагом $h = 0,5$)
0,25	0,941176	$S_{0,25} = 0, \underline{782794}$ (решение с шагом $h/2 = 0,25$)
0,50	0,800000	$\bar{S} = 0, \underline{785392}$ (решение, уточненное по методу Рунге)
0,75	0,640000	
1,00	0,500000	$S = 0,785398$ (точное решение)

При $h = 0,5$ по квадратурной формуле метода трапеций (2.21) получим

$$S_{0,5} = 0,5 \left[\frac{y(0) + y(1)}{2} + y(0,5) \right] = 0,5 \left[\frac{1 + 0,5}{2} + 0,8 \right] = 0,775.$$

При $h = 0,25$ получим

$$S_{0,25} = 0,25 \left[\frac{1 + 0,5}{2} + 0,941176 + 0,8 + 0,64 \right] = 0,782794.$$

Вычислим член, приближенно определяющий погрешность

$$R \approx \frac{S_{0,25} - S_{0,5}}{3} = 0,002598.$$

Для метода трапеций уточненное значение для $S_{0,25}$ определим по формуле (2.44)

$$\bar{S} = S_{0,25} + R = 0,782794 + 0,002598 = 0,785392.$$

Точное значение интеграла равно $S = 0,785398$ (см. пример 4 в подразделе 2.3). В табл. 2.4 приведены полученные результаты. Чертой внизу отмечены верные цифры. Видно, что до уточнения величина $S_{0,25}$ имела две верные цифры после запятой, а уточненное решение $\bar{S}$ имеет уже пять верных цифр после запятой. Легко посчитать фактическую погрешность для величин $S_{0,25}$ и $\bar{S}$:

$$S - S_{0,25} = 0,785398 - 0,782794 = 0,002604;$$

$$S - \bar{S} = 0,785398 - 0,785392 = 0,000006.$$

Видно, что уточнение по методу Рунге уменьшает в данном случае погрешность более чем в 400 раз.

Завершая рассмотрение методов численного интегрирования, отметим, что в математическом обеспечении всех ЭВМ имеются эффективные программы численного интегрирования. Эти программы обычно реализуют **адаптивные алгоритмы**, которые используют разные значения шага интегрирования на отдельных участках отрезка интегрирования, позволяя уменьшить машинное время без потери точности результатов. Это вызвано тем, что подынтегральная функция может по-разному вести себя на отдельных участках интегрирования: слабо изменяться на одних участках и сильно расти (или убывать) на других. Ясно, что на участках медленного изменения функции шаг интегрирования может быть достаточно большим, а на участках сильного изменения функции шаг приходится брать очень мелким. Адаптивные алгоритмы позволяют автоматически подбирать шаг на отдельных интервалах отрезка интегрирования и путем измельчения шага оценивать на каждом интервале погрешность (например, по методу Рунге) и выбирать шаг таким образом, чтобы получить заданную точность.

3. ЧИСЛЕННОЕ ДИФФЕРЕНЦИРОВАНИЕ

3.1. Постановка задачи

Под термином **численное дифференцирование** обычно понимается приближенное вычисление производных для таблично заданной функции. Помимо этого основного случая, численное дифференцирование используется для вычисления производной аналитически заданной функции, когда производная имеет слишком сложное для вычислений выражение.

Напомним, что производной функции $y = f(x)$ называется предел отношения приращения функции $\Delta y = f(x + \Delta x) - f(x)$ к приращению аргумента Δx при стремлении Δx к нулю

$$\frac{dy}{dx} = f'(x) = \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x}. \quad (3.1)$$

Если функция задана аналитически (в виде формулы), то для вычисления производных используют готовые формулы и правила, которые изучаются в курсе высшей математике. При численном нахождении производной можно воспользоваться формулой (3.1) без предельного перехода $\Delta x \rightarrow 0$, используя приближенное равенство

$$y' \approx \Delta y / \Delta x, \quad (3.2)$$

где Δx – конечное число, а Δy – конечная разность, вычисляемая по определенным правилам. В зависимости от способа вычисления Δy можно получить различные формулы для вычисления производной в одной и той же точке. В первом разделе уже вводилось понятие конечных разностей при получении интерполяционного многочлена Ньютона [Формулы (1.16) – (1.18)]. Пусть функция $y = f(x)$ задана таблично: множеству значений аргументов $x_0, x_1, \dots, x_n$ соответствует множество значений функции $y_0, y_1, \dots, y_n$. Введем конечные разности в точке x_1 : $\Delta y_1 = y_2 - y_1$ (правая разность), $\Delta y_1 = y_1 - y_0$ (левая разность), $\Delta y_1 = y_2 - y_0$ (центральная разность). Пусть шаг изменения аргумента $h = x_{i+1} - x_i$ в таблице является постоянным. Тогда производную y'_1 в точке $x = x_1$ можно определить с помощью приближенных соотношений:

$$y'_1 \approx (y_2 - y_1)/h; \quad (3.3)$$

$$y'_1 \approx (y_1 - y_0)/h; \quad (3.4)$$

$$y'_1 \approx (y_2 - y_0)/2h. \quad (3.5)$$

Аналогично можно вычислить вторую производную

$$y''_1 = (y'_1)' \approx \frac{y'_2 - y'_1}{h} = \frac{(y_2 - y_1)/h - (y_1 - y_0)/h}{h} = \frac{y_2 - 2y_1 + y_0}{h^2}. \quad (3.6)$$

В числителе последнего выражения (3.6) стоит центральная вторая разность $\Delta^2 y_1 = \Delta y_2 - \Delta y_1 = (y_2 - y_1) - (y_1 - y_0) = y_2 - 2y_1 + y_0$. Формулы (3.3) – (3.6) будут часто использоваться в этом и следующих разделах. Однако подход, основанный на непосредственном использовании соотношения (3.2), является недостаточно общим. Он не позволяет получать многие необходимые для практики выражения для производных и оставляет открытым вопрос о погрешности вычисления производных.

Рассмотрим более общий подход к задаче численного дифференцирования. Аппроксимируем функцию $f(x)$ некоторой аппроксимирующей функцией $\varphi(x)$

$$f(x) = \varphi(x) + R(x). \quad (3.7)$$

В качестве аппроксимирующей функции $\varphi(x)$ можно взять любую интерполяционную функцию (например, многочлен или тригонометрический многочлен), функцию наилучшего равномерного приближения или наилучшую функцию, аппроксимирующую “в среднем” функцию $f(x)$ с помощью метода наименьших квадратов. Функция $R(x)$ в (3.7) – погрешность аппроксимации. Дифференцируя выражение (3.7), найдем значение производной $f'(x)$

$$f'(x) = \varphi'(x) + R'(x). \quad (3.8)$$

В качестве приближенного значения производной $f'(x)$ можно принять соответствующее значение производной $\varphi'(x)$ интерполирующей функции:

$$f'(x) \approx \varphi'(x).$$

Функция $R'(x) = f'(x) - \varphi'(x)$, характеризующая величину отклонения приближенного значения производной от ее истинного значения, называется **погрешностью аппроксимации производной**.

Аналогично можно определить любую производную более высокого порядка

$$f^{(k)}(x) = \varphi^{(k)}(x) + R^{(k)}(x), \quad f^{(k)}(x) \approx \varphi^{(k)}(x). \quad (3.9)$$

Отметим принципиальную особенность задачи дифференцирования: она является некорректной в том смысле, что малые изменения значений функции могут привести к довольно большим различиям в значениях производных. Поясним это на примере. Пусть $f(x)$ и $\tilde{f}(x)$ – две непрерывно дифференцируемые функции на отрезке $[a, b]$, связанные соотношением

$$\tilde{f}(x) = f(x) + \frac{1}{n} \cos(n^2 x). \quad (3.10)$$

Разность значений этих функций ограничена и при больших n может быть сделана сколь угодно малой. Действительно,

$$\left| \tilde{f}(x) - f(x) \right| = \left| \frac{1}{n} \cos(n^2 x) \right| \leq \frac{1}{n} \rightarrow 0 \quad (n \rightarrow \infty).$$

Это неравенство следует из ограниченности функции косинуса при любом x : $|\cos(n^2 x)| \leq 1$. Возьмем теперь производную от обеих частей равенства (3.10)

$$\tilde{f}'(x) = f'(x) - \frac{n^2}{n} \sin(n^2 x) = f'(x) - n \sin(n^2 x),$$

откуда следует, что

$$\left| \tilde{f}'(x) - f'(x) \right| = n |\sin(n^2 x)|. \quad (3.10a)$$

Из последнего равенства следует, что для некоторых значений x разность производных $f'(x)$ и $\tilde{f}'(x)$ может оказаться равной n и, следовательно, может быть сколь угодно большой. Это фактически означает, что если мы аппроксимируем функцию $f(x)$ с помощью аппроксимирующей функции $\varphi(x)$ и имеют место равенства (3.7) и (3.8), то из малости погрешности аппроксимации не всегда следует малость погрешности аппроксимации производной $R'(x) = f'(x) - \varphi'(x)$. Это надо иметь в виду при практическом применении численного дифференцирования, особенно в случае, когда исходная функция $f(x)$ определяется экспериментально и, следовательно, величины y_i заданы с некоторой ошибкой. Для того чтобы эти ошибки не привели к значительным погрешностям в определении производных, применяются специальные приемы сглаживания исходной экспериментальной функции. Некоторые из этих приемов описаны в работе [7].

Рассмотрим различные способы получения формул численного дифференцирования.

3.2. Использование ряда Тейлора

Пусть функция $y = f(x)$ задана с помощью таблицы на отрезке $[a, b]$. Обозначим множество ее аргументов через $x_0, x_1, \dots, x_n$ ($x_0 = a, x_n = b$), а множество значений функции через $y_0, y_1, \dots, y_n$. Будем считать шаг изменения аргумента постоянным: $x_{i+1} - x_i = x_i - x_{i-1} = \text{const} = h$. Рассмотрим произвольную внутреннюю точку x_i и две соседних с ней точки x_{i+1} и x_{i-1} . Значения функции

в этих точках обозначим через y_i, y_{i+1}, y_{i-1} : $y_{i-1} = f(x_{i-1})$, $y_i = f(x_i)$, $y_{i+1} = f(x_{i+1})$. Будем считать функцию достаточно гладкой и разложим ее в ряд Тейлора в окрестности базовой точки x_i . С помощью этого ряда значения функции $y = f(x)$ в соседних справа и слева точках x_{i+1} и x_{i-1} можно записать в виде:

$$y_{i+1} = y_i + y'_i h + \frac{y''_i h^2}{2!} + \frac{y'''_i h^3}{3!} + \frac{y^{(4)}_i h^4}{4!} + O(h^5); \quad (3.11)$$

$$y_{i-1} = y_i - y'_i h + \frac{y''_i h^2}{2!} - \frac{y'''_i h^3}{3!} + \frac{y^{(4)}_i h^4}{4!} - O(h^5). \quad (3.12)$$

Здесь использованы обозначения: $y'_i = y'(x_i)$, $y''_i = y''(x_i)$, $y'''_i = y'''(x_i)$. Переносим первый член суммы в правой части (3.11) в левую часть и разделив обе части на h , получим:

$$\frac{y_{i+1} - y_i}{h} = y'_i + \frac{y''_i}{2!} h + O(h^2) = y'_i + O(h); \quad (3.13)$$

Аналогично из (3.12) следует:

$$\frac{y_i - y_{i-1}}{h} = y'_i - \frac{y''_i}{2!} h + O(h^2) = y'_i - O(h). \quad (3.14)$$

Считая h малым и пренебрегая членами $O(h)$, получим конечно-разностные формулы для приближенной аппроксимации первой производной с помощью правой и левой разностей:

$$\nabla y'_i = \frac{y_{i+1} - y_i}{h}; \quad (3.15)$$

$$\Delta y'_i = \frac{y_i - y_{i-1}}{h}. \quad (3.16)$$

Вычитая почленно из (3.11) члены равенства (3.12) и разделив обе части полученного равенства на $2h$, получим

$$\frac{y_{i+1} - y_{i-1}}{2h} = y'_i + \frac{y''_i h^2}{3!} + O(h^4) = y'_i + O(h^2). \quad (3.17)$$

Пренебрегая членом второго порядка малости $O(h^2)$, получим более точную аппроксимацию для первой производной, использующую центральную разность

$$y'_i = \frac{y_{i+1} - y_{i-1}}{2h}. \quad (3.18)$$

Соотношения (3.13), (3.14) и (3.17) позволяют получить выражения для погрешности аппроксимации первой производной по формулам (3.15), (3.16) и (3.18). Эти погрешности равны: $y''_i h/2$ для (3.15), $-y''_i h/2$ для (3.16), $y'''_i h^2/6$ для (3.18). Видно, что погрешность зависит от значения некоторой производной в базовой точке x_i и шага таблицы h .

Зависимость от шага обычно записывают в виде $O(h^k)$. Показатель степени k называют **порядком погрешности аппроксимации производной** (или просто порядком аппроксимации). Равенства (3.13) и (3.14) показывают, что правые и левые разности дают первый порядок аппроксимации ($k = 1$), а центральная разность дает второй порядок ($k = 2$), что обеспечивает более высокую точность. Полученные ранее формулы (3.3), (3.4) и (3.5) есть частный случай общих формул (3.15), (3.16) и (3.18) при $i = 1$. Очевидно, что для аппроксимации первой производной y'_i лучше применять более точную формулу (3.18), использующую центральные разности. Формулы (3.15) и (3.16) с односторонними разностями можно использовать в крайних точках интервала $[a, b]$ при $x_0 = a$ или $x_n = b$.

Получим теперь выражение для аппроксимации второй производной y''_i . Для этого сложим левые и правые части выражений (3.11) и (3.12). Члены с разными знаками при этом сократятся

$$y_{i+1} + y_{i-1} = 2y_i + y''_i h^2 + \frac{y^{(4)}_i h^4}{12} + O(h^6).$$

Переносим член $2y_i$ в левую часть и разделив обе части на h^2 , получим, отбросив малую более высокого порядка $O(h^6)$,

$$y''_i = \frac{y_{i+1} - 2y_i + y_{i-1}}{h^2} - \frac{y^{(4)}_i h^2}{12}. \quad (3.19)$$

В формуле (3.19) в явном виде выписано выражение для погрешности аппроксимации (последний член в правой части). Видно, что эта аппроксимация имеет второй порядок точности по h и зависит от четвертой производной функции $y = f(x)$ в базовой точке x_i . Считая эту погрешность малой, получим

$$y_i'' = (y_{i+1} - 2y_i + y_{i-1})/h^2. \quad (3.20)$$

Приведем пример использования полученных формул.

Пример 8. Задана табличная функция $y = f(x)$

x	0,2	0,3	0,4
y	0,008	0,027	0,064

Вычислить значения $y_i'(x)$ при всех трех значениях аргументов и значение $y_i''(x)$ в точке $x = 0,3$.

Решение. Заметим прежде всего, что в таблице приведена функция $y = x^3$, что позволяет, сравнив с точным решением $y' = 3x^2$, $y'' = 6x$, $y''' = 6$, $y^{(4)} = 0$, оценить ошибку аппроксимации. В данном случае $h = 0,1$.

Для вычисления $y'(0,2)$ используем (3.15):

$$y'(0,2) = (0,027 - 0,008)/0,1 = 0,019.$$

Точное значение $y'(0,2) = 3 \cdot (0,2)^2 = 0,12$. Далее везде рядом с приближенным решением будем в квадратных скобках указывать точное значение соответствующей производной.

Для вычисления $y'(0,4)$ используем формулу (3.16):

$$y'(0,4) = (0,064 - 0,027)/0,1 = 0,37 \quad [y'(0,4) = 3 \cdot (0,4)^2 = 0,48].$$

Для вычисления $y'(0,3)$ используем более точную формулу (3.18)

$$y'(0,3) = (0,064 - 0,008)/(2 \cdot 0,1) = 0,28 \quad [y'(0,3) = 3 \cdot (0,3)^2 = 0,27].$$

Для второй производной $y''(0,3)$ используем (3.20)

$$y''(0,3) = (0,064 - 2 \cdot 0,027 + 0,008)/(0,1)^2 = 1,8 \quad [y''(0,3) = 6 \cdot 0,3 = 1,8].$$

Таблица 3.1

x_i	y'_i вычисленное	y'_i точный рез-тат	Фактическая погрешность	Погрешность аппроксимации
0,2	0,19	0,12	0,07	$(y'_i h)/2 = 0,06$
0,3	0,28	0,27	0,01	$(y'_i h)/6 = 0,01$
0,4	0,37	0,48	-0,11	$-(y'_i h)/2 = -0,12$
x_i	y''_i вычисленное	y''_i точный рез-тат	Фактическая погрешность	Погрешность аппроксимации
0,3	1,8	1,8	0	$-(y''_i h^2)/12 = 0$

Для удобства сравнения приведем полученные результаты в табл. 3.1.

Сделаем некоторые замечания относительно результатов, приведенных в табл. 3.1. Во-первых, видно, что вычисление $y'(x)$ с помощью центральной разности [формула (3.18)] дает значительно более точные результаты, чем односторонние разности [формулы (3.15) и (3.16)], так как у последних погрешность имеет первый порядок $O(h)$. Во-вторых, сравнение фактической погрешности аппроксимации вычисления производных показывает, что они отличаются незначительно. Небольшое отличие объясняется влиянием членов более высокого порядка малости, входящих в разложение функции в ряд Тейлора. В-третьих, полное совпадение вычисляемого и точного значения величины $y''(0,3)$ объясняется тем, что погрешность аппроксимации для рассматриваемой функции $y = x^3$ равна нулю, так как входящая в формулу погрешности аппроксимации величина $f^{(4)}(x)$ и все производные более высокого порядка равны нулю [см. формулу (3.19)].

3.3. Использование интерполяционных многочленов

Остановимся на другом способе получения формул численного дифференцирования, основанном на использовании интерполяционного многочлена Лагранжа. Для удобства рассмотрим равномерное расположение точек x_i с шагом $x_i - x_{i-1} = \text{const} = h$, хотя этот способ можно использовать для получения формул аппроксимации производных с неравномерным расположением аргументов функции $y = f(x)$.

Рассмотрим опять три рядом расположенные точки x_{i-1} , x_i и x_{i+1} и соответствующие значения функции y_{i-1}, y_i, y_{i+1} . Запишем интерполя-

ционный многочлен Лагранжа $L(x_i)$ для случая трех узлов интерполяции x_{i-1} , x_i , x_{i+1} [формула (1.15)] и соответствующий ему остаточный член $R_L(x)$, определяющий погрешность интерполяции:

$$L(x) = \frac{1}{2h^2} [(x-x_i)(x-x_{i+1})y_{i-1} - 2(x-x_{i-1})(x-x_{i+1})y_i + (x-x_{i-1})(x-x_i)y_{i+1}], \quad (3.21)$$

$$R_L(x) = \frac{y'''(\zeta)}{3!} (x-x_{i-1})(x-x_i)(x-x_{i+1}), \quad (\zeta \in [x_{i-1}, x_{i+1}]). \quad (3.22)$$

Непосредственной подстановкой легко убедиться, что $L(x_{i-1}) = y_{i-1}$, $L(x_i) = y_i$, $L(x_{i+1}) = y_{i+1}$, $R_L(x_{i-1}) = R_L(x_i) = R_L(x_{i+1}) = 0$ и, следовательно, $L(x)$ действительно является квадратным трехчленом, проходящим через три точки (x_{i-1}, y_{i-1}) , (x_i, y_i) и (x_{i+1}, y_{i+1}) :

$$f(x) = L(x) + R_L(x). \quad (3.23)$$

Положим теперь $f'(x) \approx L'(x)$, причем погрешность аппроксимации производной легко найти, дифференцируя обе части (3.23)

$$f'(x) = L'(x) + R'_L(x). \quad (3.24)$$

Непосредственное дифференцирование $L(x)$ и $R_L(x)$ дает следующие соотношения:

$$L'(x) = \frac{1}{2h^2} [(2x-x_i-x_{i+1})y_{i-1} - 2(2x-x_{i-1}-x_{i+1})y_i + (2x-x_{i-1}-x_i)y_{i+1}]; \quad (3.25)$$

$$R'_L(x) = \frac{y'''(\zeta)}{3!} [(x-x_i)(x-x_{i+1}) + (x-x_{i-1})(x-x_{i+1}) + (x-x_{i-1})(x-x_i)], \quad (3.26)$$

где $y'''(\zeta)$ – значение производной третьего порядка в некоторой внутренней точке $(\zeta \in [x_{i-1}, x_{i+1}])$. Формула (3.25) для $L'(x)$ дает возможность приближенно вычислить $f'(x)$ для исходной функции, причем получаемая при этом погрешность определяется по формуле (3.26) с точностью до неизвестной, вообще говоря, величины $y'''(\zeta)$.

Воспользуемся формулой (3.25) для получения конечно-разностных формул для аппроксимации производной $f'(x)$ в узловых точках x_{i-1} ,

x_i, x_{i+1} . Для $x = x_{i-1}$ получим

$$\begin{aligned} y'_{i-1} &= y'(x_{i-1}) = L'(x_{i-1}) + R'_L(x_{i-1}) = \\ &= \frac{1}{2h^2} [(2x_{i-1} - x_i - x_{i+1})y_{i-1} - 2(2x_{i-1} - x_{i-1} - x_{i+1})y_i + (2x_{i-1} - x_{i-1} - x_i)y_{i+1}] + \\ &+ \frac{y'''(\zeta)}{3!} [(x_{i-1} - x_i)(x_{i-1} - x_{i+1}) + (x_{i-1} - x_{i-1})(x_{i-1} - x_{i+1}) + (x_{i-1} - x_{i-1})(x_{i-1} - x_i)]. \end{aligned}$$

Подставляя в это выражение величину $h = x_{i+1} - x_i = x_i - x_{i-1}$ и отбрасывая члены, равные нулю, получим

$$y'_{i-1} = \frac{1}{2h} (-3y_{i-1} + 4y_i - y_{i+1}) + \frac{h^2}{3} y'''(\zeta). \quad (3.27)$$

Аналогичное соотношение для точки $x = x_i$ имеет вид

$$y'_i = \frac{1}{2h} (y_{i+1} - y_{i-1}) - \frac{h^2}{6} y'''(\zeta). \quad (3.28)$$

Для точки $x = x_{i+1}$ получим

$$y'_{i+1} = \frac{1}{2h} (y_{i-1} - 4y_i + 3y_{i+1}) + \frac{h^2}{3!} y'''(\zeta). \quad (3.29)$$

Формула (3.28) – это уже знакомая нам аппроксимация первой производной с помощью центральной разности, которая совпадает с (3.17) предыдущего подраздела. Формулы (3.27) и (3.29) дают возможность вычислять первую производную в крайних точках со вторым порядком точности, используя значения функции в трех точках y_{i-1}, y_i и y_{i+1} . Формулы с односторонними разностями (3.15) и (3.16) дают погрешность первого порядка и используют лишь две точки: y_{i-1} и y_i для y'_{i-1} (или y_i и y_{i+1} для y'_{i+1}).

Еще раз продифференцировав (3.24), получим

$$f''(x) = L''(x) + R''_L(x). \quad (3.30)$$

Вычисляя производные от (3.25) и (3.26), можно получить конкретные формулы для $L''(x)$ и $R_L''(x)$, а с их помощью – выражение для $f''(x)$ в любой точке $x \in [x_{i-1}, x_{i+1}]$ и выражение для погрешности аппроксимации $R_L''(x)$. Не проводя всех выкладок, приведем окончательные конечно-разностные формулы для приближенного вычисления второй производной $y'' = f''(x)$ в узловых точках x_{i-1} , x_i , x_{i+1} , использующие три узла интерполяции:

$$y''_{i-1} = \frac{1}{h^2}(y_{i-1} - 2y_i + y_{i+1}) + O(h); \quad (3.31)$$

$$y''_i = \frac{1}{h^2}(y_{i-1} - 2y_i + y_{i+1}) + O(h^2); \quad (3.32)$$

$$y''_{i+1} = \frac{1}{h^2}(y_{i-1} - 2y_i + y_{i+1}) + O(h). \quad (3.33)$$

Эти формулы имеют одинаковый вид, но первая и третья имеют первый порядок точности, а вторая – второй порядок точности. Обратим внимание, что в (3.32) стоит центральная разность, когда производная вычисляется в центральной точке $x = x_i$, относительно которой выражение для конечной разности является симметричным. Сравнение выражений (3.13), (3.14) и (3.17) показывает, что аналогично обстоит дело и для первой производной: использование центральной разности дает на порядок более высокую точность, чем несимметричные разности. Отсюда следует, что использование центральных разностей, симметричных относительно точки, в которой вычисляется производная, более выгодно по сравнению с другими конечно-разностными выражениями с точки зрения получаемой точности.

Выше мы рассматривали интерполяционный многочлен второй степени, проходящий через три узла, и получили с его помощью конечно-разностные формулы для первой и второй производных в узловых точках. Аналогично можно использовать четыре, пять и большее количество узлов, записывая соответственно многочлены $L(x)$ третьей, четвертой и более высоких степеней. В этом случае мы можем получить конечно-разностные выражения для y' , y'' , y''' и т.д. с более высокой точностью, но требующие привлечения большего количества узлов. В качестве примера приведем без вывода две формулы, использующие центральные разности для получения $y'(x)$ и $y''(x)$ в центральном

узле с использованием пяти узлов интерполяции с помощью многочлена четвертой степени:

$$y'_i = \frac{1}{12h}(y_{i-2} - 8y_{i-1} + 8y_{i+1} - y_{i+2}) + O(h^4), \quad (3.34)$$

$$y''_i = \frac{1}{12h^2}(-y_{i-2} - 16y_{i-1} - 30y_i + 16y_{i+1} - y_{i+2}) + O(h^4). \quad (3.35)$$

3.4. Погрешности численного дифференцирования и способы их уменьшения

В предыдущих подразделах мы рассмотрели лишь один из источников погрешности численного дифференцирования – погрешность аппроксимации. Однако существуют еще два вида погрешности, возникающие при численном дифференцировании. Во-первых, это возможная неточность задания значений функции в узлах. Эта неточность может быть вызвана как ошибками при получении экспериментальных данных, так и вычислительной погрешностью, если табличная функция $y = f(x)$ получена с помощью некоторых вычислительных процедур. Во-вторых, это погрешности округления при реализации методов численного дифференцирования. Вообще говоря, погрешность аппроксимации уменьшается при уменьшении шага h , так как она обычно имеет вид ch^k , где k – порядок погрешности. Напротив, погрешности округления возрастают с уменьшением шага h . Поэтому суммарная погрешность численного дифференцирования может убывать при уменьшении шага лишь до некоторого предельного значения h_* , после чего дальнейшее уменьшение шага не только не повысит точности результатов, а, наоборот, может привести к росту суммарной погрешности. Способы определения примерной величины h_* с учетом погрешностей аппроксимации, округления и неточности задания исходных значений функции y_i приведены, например, в работе [3].

Если $h > h_*$, то, так же как и в численном интегрировании, существует простой и эффективный способ уточнения аппроксимации производной по результатам расчетов с различными шагами аргумента функции при заданном порядке аппроксимации – **метод Рунге**. Изложим суть этого метода. Пусть $f'(x)$ вычисляется с помощью какой-либо конечно-разностной аппроксимации на равномерной сетке с шагом h . Тогда можно записать, что

$$f'(x) = M(x, h) + R. \quad (3.36)$$

Здесь $M(x, h)$ – конечно-разностная аппроксимация для производной $f'(x)$, где в явном виде указана зависимость этой аппроксимации от h , а R – погрешность аппроксимации, которую, как мы видели, можно записать в виде

$$R = h^k \varphi(x) + O(h^{k+1}), \quad (3.37)$$

где k – порядок аппроксимации производной. Значит, формулу (3.36) можно записать в виде

$$f'(x) = M(x, h) + h^k \varphi(x) + O(h^{k+1}). \quad (3.38)$$

Запишем теперь соотношение (3.38) в той же точке x с помощью той же конечно-разностной аппроксимации, но с вдвое большим шагом $2h$

$$f'(x) = M(x, 2h) + (2h)^k \varphi(x) + O((2h)^{k+1}). \quad (3.39)$$

Так как в (3.38) и (3.39) слева стоит одна и та же величина, то приравнявая их правые части, можно найти выражение для главного члена погрешности аппроксимации производной, если пренебречь малыми более высокого порядка

$$h^k \varphi(x) = \frac{M(x, h) - M(x, 2h)}{2^k - 1}. \quad (3.40)$$

Формулу (3.38) теперь можно записать в виде

$$f'(x) = M(x, h) + \frac{M(x, h) - M(x, 2h)}{2^k - 1} + O(h^{k+1}). \quad (3.41)$$

Эта формула позволяет по результатам двух расчетов значений производной в точке x $M(x, h)$ и $M(x, 2h)$ (с шагами h и $2h$) при заданном порядке аппроксимации k найти ее уточненное значение с порядком точности $k + 1$. Приведем пример использования правила Рунге.

Пример 9. С помощью метода Рунге найти численно производную $y'(1)$ для таблично заданной функции

x	0,8	0,9	1,0
y	0,512	0,729	1,0

Решение. Вычислим величину $y'(1)$ с помощью левой разности [формула (3.14)],

$$y'_i = \frac{y_i - y_{i-1}}{h} + O(h),$$

порядок погрешности которой $k = 1$. Воспользуемся этой формулой, беря шаги $h = 0,1$ и $2h = 0,2$. При $h = 0,1$

$$y'_h(1) = \frac{1 - 0,729}{0,1} = 2,71.$$

При $h = 0,2$

$$y'_{2h}(1) = \frac{1 - 0,512}{0,2} = 2,44.$$

Вычислим уточненное значение производной $y'(1)$ по формуле

$$y'(1) = y'_h + \frac{y'_h - y'_{2h}}{2^1 - 1} = 2,71 + \frac{2,71 - 2,44}{2 - 1} = 2,71 + 0,27 = 2,98.$$

Для оценки укажем, что в таблице приведена функция $y = x^3$, точное значение производной которой равно $y' = 3x^2$ и, следовательно, $y'(1) = 3$. Видно, что уточненное с помощью метода Рунге значение $y'(1) = 2,98$ имеет гораздо меньшую ошибку, чем значения $y'_h = 2,71$ и $y'_{2h} = 2,44$.

4. РАЗНОСТНЫЕ УРАВНЕНИЯ

4.1. Общие понятия

Многие теоретические и прикладные задачи различных разделов науки, техники и экономики приводят к дифференциальным уравнениям. В некоторых случаях удается получить решение этих уравнений аналитическим путем, выразив его в виде формулы, зависящей от элементарных функций. Методы получения таких решений описаны в курсе высшей математики. Однако в большинстве случаев получить аналитическое

решение невозможно и приходится использовать численное решение этих задач. Одним из наиболее универсальных методов численного решения дифференциальных уравнений является метод конечных разностей. Сущность этого метода заключается в том, что функцию $y(x)$, являющуюся решением дифференциального уравнения, мы заменяем дискретной функцией, определенной на некотором множестве дискретных аргументов, называемом обычно сеткой. **Дискретная (сеточная) функция** – это функция, в которой дискретному множеству значений аргумента, состоящему из узлов сетки, ставится в соответствие дискретное множество значений функции. Приближенные методы, в которых функции непрерывного аргумента, являющиеся решениями дифференциальных или интегральных уравнений, заменяются сеточными функциями, называются **сеточными методами**. Получающиеся при этом уравнения, связывающие неизвестные значения сеточной функции при нескольких значениях дискретного аргумента, называются **разностными уравнениями**.

В первом разделе мы уже рассматривали понятие конечных разностей при получении интерполяционного многочлена Ньютона. Пусть задана функция $y = f(x)$, которая определена в точках $x_0, x_1 = x_0 + h, x_2 = x_0 + 2h, \dots, x_n = x_0 + nh$ и принимает в них значения $y_0 = f(x_0), y_1 = f(x_1), \dots, y_n = f(x_n)$. Напомним, что конечными разностями первого порядка называют разность двух последовательных значений функции $\Delta y_0 = y_1 - y_0, \Delta y_1 = y_2 - y_1, \dots, \Delta y_{n-1} = y_n - y_{n-1}$. Разности между двумя последовательными разностями первого порядка называют разностями второго порядка или просто вторыми разностями:

$$\Delta^2 y_0 = \Delta y_1 - \Delta y_0 = y_2 - 2y_1 + y_0,$$

$$\Delta^2 y_1 = \Delta y_2 - \Delta y_1 = y_3 - 2y_2 + y_1,$$

$$\Delta^2 y_i = \Delta y_{i+1} - \Delta y_i = y_{i+2} - 2y_{i+1} + y_i, \dots$$

При записи вторых разностей показано, что они в явном виде могут быть выражены через значения функции y_i ($i = 0, 1, \dots, n$). Не приводя подробно всех формул, заметим, что аналогично могут быть введены разности более высоких порядков. Разность m -го порядка определяется как



$$\Delta^m y_i = \Delta^{m-1} y_{i+1} - \Delta^{m-1} y_i \quad (i = 0, 1, 2, \dots). \quad (4.1)$$

Точно так же, как это было сделано для разностей второго порядка, разности любого порядка можно выразить через значения функции y_i в соответствующих точках дискретного аргумента.

Иногда используется другое определение разностного уравнения, эквивалентное, вообще говоря, тому определению, которое было дано выше. Разностным называют уравнение, связывающее аргумент x , неизвестную функцию $y = f(x)$ и ее конечные разности $\Delta y, \Delta^2 y, \dots, \Delta^m y$. Порядком разностного уравнения в этом случае называется наивысший порядок входящей в уравнение разности. Разностное уравнение m -го порядка в общем случае имеет вид

$$P(x, y(x), \Delta y(x), \Delta^2 y(x), \dots, \Delta^m y(x)) = 0. \quad (4.2)$$

Если вместо конечных разностей в уравнении (4.2) подставить их выражение через функцию, то получим другую форму разностного уравнения

$$P_1(x, y(x), y(x+h), y(x+2h), \dots, y(x+nh)) = 0. \quad (4.3)$$

Мы будем использовать разностные уравнения в форме (4.3). В этом случае порядок разностного уравнения определяется количеством значений функции в различных подряд расположенных точках, которые входят в разностное уравнение. Приведем примеры разностных уравнений:

$$y(x) + 2y(x+h) = \varphi(x); \quad (4.4)$$

$$3y(x-h) + 5y(x) + y(x+h) = q(x); \quad (4.5)$$

$$y(x-h) + y(x+h) = \psi(x). \quad (4.6)$$

Уравнение (4.4), куда входят значения функции в двух точках x и $x+h$, называют разностным уравнением первого порядка. Уравнение, в которое входят значения функции в трех соседних точках, называют разностным уравнением второго порядка. Уравнение (4.6) также будет уравнением второго порядка, хотя в него входят функции только в двух точках. В данном случае можно считать, что коэффициент при $y(x)$ равен нулю. Максимальный порядок разности аргументов неизвестной функции $y(x)$ определяет порядок разностного уравнения. В этом смысле и уравнение (4.5), и уравнение (4.6) имеют второй порядок, так как максимальная разность аргументов у обоих уравнений равна $x+h - (x-h) = 2h$.

Теория разностных уравнений весьма тесно связана с теорией обыкновенных дифференциальных уравнений. Подходы и методы их решения имеют много общего, особенно если речь идет об уравнениях с постоянными коэффициентами. Вопрос об их решении мы рассмотрим позже, а сейчас приведем простой пример сведения дифференциального уравнения к разностному.

Рассмотрим дифференциальное уравнение

$$y'(x) + Ay(x) = 0, \quad (4.7)$$

удовлетворяющее начальному условию $y(0) = 1$. Решением этого уравнения, удовлетворяющим начальному условию, будет функция $y = e^{-Ax}$. Это легко проверить непосредственным дифференцированием $y(x)$ и подстановкой в уравнение (4.7). Будем решать эту задачу численно на отрезке $[0, 1]$. Разобьем этот отрезок на n одинаковых интервалов, длина каждого из которых равна h ($h = 1/n$). Вместо функции $y(x)$ будем искать табличную функцию, заданную на дискретном множестве аргументов: $y(0), y(h), y(2h), \dots, y(nh)$. Заменяем производную $y'(x)$ разностным отношением

$$\frac{y(x+h) - y(x)}{h} + Ay(x) = 0. \quad (4.8)$$

В результате вместо дифференциального уравнения (4.7) мы получаем приближающее его разностное уравнение (4.8), которое перепишем в более удобном для нас виде, умножив обе части уравнения на h и приведя подобные члены

$$y(x+h) = (1 - Ah) y(x). \quad (4.9)$$

Уравнение (4.9) [как и уравнение (4.8)] является разностным уравнением первого порядка, решение которого в данном случае можно легко найти, используя начальное условие $y(0) = 1$. Полагая последовательно $x = 0, h, 2h, \dots, nh$, получим:

$$\begin{aligned} y(h) &= (1 - Ah)y(0) = (1 - Ah); \\ y(2h) &= (1 - Ah)y(h) = (1 - Ah)^2; \\ y(3h) &= (1 - Ah)y(2h) = (1 - Ah)^3; \\ &\dots\dots\dots \\ y(nh) &= (1 - Ah)y((n-1)h) = (1 - Ah)^n. \end{aligned}$$

Так как $nh = 1$, $h = 1/n$, то из последнего соотношения получим

$$y(1) = \left(1 - \frac{A}{n}\right)^n. \quad (4.10)$$

Решение дифференциального уравнения (4.7) $y = e^{-Ax}$ при $x = 1$ дает $y = e^{-A}$. Используя второй замечательный предел (типа e), легко показать, что

$$\lim_{n \rightarrow \infty} \left(1 - \frac{A}{n}\right)^n = e^{-A}.$$

Отсюда следует, что приближенное решение, получаемое по разностной схеме (4.8) и зависящее от h , при уменьшении шага h (и увеличении, как следствие, величины n) сходится к точному решению дифференциального уравнения.

Рассмотрим теперь дифференциальное уравнение второго порядка

$$y''(x) + Ay'(x) + By(x) = f(x). \quad (4.11)$$

Заменяя $y''(x)$ и $y'(x)$ конечно-разностными соотношениями, использующими центральные разности и описанными в третьем разделе, получим разностное уравнение

$$\frac{y(x+h) - 2y(x) + y(x-h))}{h^2} + A \frac{y(x+h) - y(x-h)}{2h} + By(x) = f(x). \quad (4.12)$$

Приведя подобные члены, его можно записать в виде

$$\frac{1}{h^2} \left(1 - \frac{Ah}{2}\right) y(x-h) - \frac{1}{h^2} (2 - Bh^2) y(x) + \frac{1}{h^2} \left(1 + \frac{Ah}{2}\right) y(x+h) = f(x). \quad (4.13)$$

Это разностное уравнение второго порядка. Для получения единственного решения дифференциального уравнения (4.11) надо задать два дополнительных условия. Аналогично для разностного уравнения (4.13) также необходимо задать два дополнительных условия.

Рассмотренные разностные уравнения (4.9) и (4.13) принадлежат к одному из двух видов разностных уравнений:

$$ay(x) + by(x+h) = f(x); \quad (4.14)$$

$$ay(x-h) + by(x) + cy(x+h) = f(x). \quad (4.15)$$

Здесь (4.14) – уравнение первого порядка, а (4.15) – уравнение второго порядка.

Способ записи разностных уравнений может быть различным. Иногда используют запись, при которой шаг изменения аргумента полагают равным единице. Этого можно достичь заменой x на hx . Тогда общая форма (4.3) разностного уравнения примет вид

$$W(x, y(x), y(x+1), y(x+1), y(x+2), \dots, y(x+n)) = 0. \quad (4.16)$$

Этот подход удобен в том случае, если не рассматривать связь между разностным уравнением и породившим его дифференциальным уравнением, а рассматривать разностные уравнения как самостоятельный объект изучения. Такой подход правомерен еще и потому, что в ряде практических приложений разностные уравнения являются самостоятельной математической моделью, не связанной с дифференциальными уравнениями. Запись разностных уравнений можно еще более упростить. Для этого последовательность точек $x_0, x_1, x_2, \dots, x_n$, делящих заданный интервал, на котором определено разностное уравнение, на отрезки длины h , занумеруем слева направо так, чтобы $x_n = x_{n-1} + h$. Также обозначим $y(x_n)$ через y_n . Тогда (4.16) примет вид

$$W(x, y_0, y_1, y_2, \dots, y_n) = 0. \quad (4.17)$$

Уравнения (4.14) и (4.15) можно записать теперь в виде:

$$ay_n + by_{n+1} = f_n; \quad (4.18)$$

$$ay_{n-1} + by_n + cy_{n+1} = f_n. \quad (4.19)$$

Здесь введено обозначение $f_n = f(x_n)$. Если считать, что разностные уравнения определены не на конечном множестве узлов, а на всей числовой оси, то соотношения (4.18) и (4.19) определяют бесконечную последовательность $\{y_n\} - \infty < n < +\infty$. На практике приходится иметь дело как с конечными, так и с бесконечными множествами сеточных функций $\{y_n\}$. Уравнения (4.18), (4.19) и другие подобные разностные уравнения более высокого порядка называются уравнениями с постоянными коэффициентами, если a, b, c не зависят от n . Например, уравнение

$$2y_n + \cos(2n+1)y_{n+1} = f_n$$

не будет уравнением с постоянными коэффициентами, так как коэффициент при y_{n+1} зависит от n .

К понятию порядка разностного уравнения можно подойти несколько иначе, чем это определялось выше. **Порядком разностного уравнения** можно назвать количество последовательных точек, задание решения в которых позволяет выделить единственное решение разностного уравнения.

Для выделения единственного решения (4.18) достаточно задать значение y в одной целой точке m , т.е. задать y_m . Действительно, уравнение (4.18) можно записать в виде рекуррентной формулы

$$y_{n+1} = \frac{1}{b}(f_n - ay_n), \quad (4.20)$$

из которой при $n = m, m+1 \dots$ последовательно определяются $y_{m+1}, y_{m+2}, \dots$, т.е. все y_n при $n > m$. Записывая (4.18) в виде другой рекуррентной формулы

$$y_{n-1} = \frac{1}{a}(f_n - by_n), \quad (4.21)$$

мы таким же путем определим все y_n при $n < m$. Таким образом, для (4.18) задание решения в одной точке позволяет выделить единственное решение для всех остальных точек. Следовательно, уравнение (4.18) имеет первый порядок.

Для выделения единственного решения уравнения (4.19) достаточно задать значения y в двух последовательных произвольно выбранных целых точках, например, задать значения y_{m-1} и y_m . Действительно, (4.19) можно записать в виде двух следующих рекуррентных формул:

$$y_{n+1} = \frac{1}{c}(f_n - by_n - ay_{n-1}), \quad (4.22)$$

$$y_{n-1} = \frac{1}{a}(f_n - by_n - cy_{n+1}). \quad (4.23)$$

По первой формуле, имея значения y_{m-1} и y_m , можно последова-

тельно определить все y_n при $n > m$, а по второй – все y_n при $n < m$. Следовательно, уравнение (4.19) – разностное уравнение второго порядка.

Как правило, разностные уравнения решаются численно. Однако есть классы сравнительно простых уравнений, для которых можно получить аналитическое решение. Приведем решение некоторых из них. Подробное описание таких решений и исследование свойств разностных уравнений содержатся в работах [5,8].

4.2. Решение разностных уравнений первого порядка

Рассмотрим разностное уравнение первого порядка вида

$$y_{n+1} = \varphi_n y_n. \quad (4.24)$$

Здесь y_n – неизвестная сеточная функция, определенная на множестве целых чисел $n = 0, 1, 2, \dots$, а φ_n – известная функция, определенная на том же множестве. Задавая в (4.24) последовательно значения независимой переменной n , равные $0, 1, 2, \dots, n-1$, получим

$$y_1 = \varphi_0 y_0, y_2 = \varphi_1 y_1, y_3 = \varphi_2 y_2, \dots, y_n = \varphi_{n-1} y_{n-1}.$$

Перемножим почленно все написанные равенства. После сокращения на произведение $y_1 y_2 y_3 \dots y_{n-1}$ получим функцию

$$y_n = [\varphi_0 \varphi_1 \varphi_2 \dots \varphi_{n-1}] y_0.$$

Вводя используемый в математике специальный символ Π , обозначающий произведение любого количества членов, получим запись этой функции в компактном виде

$$y_n = y_0 \left(\prod_{i=0}^{n-1} \varphi_i \right). \quad (4.25)$$

Величина y_0 есть начальное значение функции y_n и является произвольной постоянной: $y_0 = a_0$. Функция y_n , вычисляемая по (4.25), есть общее решение разностного уравнения (4.24).

Пример 10. Найти решение разностного уравнения

$$y_{n+1} = \frac{1}{2^n} y_n.$$

Решение. В нашем случае $\varphi_n = \frac{1}{2^n}$, поэтому решение имеет вид

$$y_n = y_0 \left(\prod_{i=0}^{n-1} \frac{1}{2^i} \right).$$

Величину произведения в данном случае легко вычислить

$$\prod_{i=0}^{n-1} \frac{1}{2^i} = 1 \cdot \frac{1}{2} \cdot \frac{1}{2^2} \cdots \frac{1}{2^{n-1}} = \frac{1}{2^{1+2+3+\dots+(n-1)}}.$$

Если воспользоваться формулой для суммы членов арифметической прогрессии, то получим

$$1 + 2 + 3 + \dots + (n-1) = \frac{n(n-1)}{2}.$$

Поэтому решение примет вид

$$y_n = y_0 \frac{1}{2^{n(n-1)/2}}. \quad (4.26)$$

Выражение (4.26) дает формулу общего решения уравнения (4.24). Рассмотрим теперь другой вид простого разностного уравнения

$$y_{n+1} = y_n + \psi_n. \quad (4.27)$$

Придавая различные целые значения переменной n , получим цепочку равенств

$$y_1 - y_0 = \psi_0, y_2 - y_1 = \psi_1, y_3 - y_2 = \psi_2, \dots, y_n - y_{n-1} = \psi_{n-1}.$$

Складывая все эти равенства, найдем решение уравнения (4.27) в виде

$$y_n = y_0 + \sum_{i=0}^{n-1} \psi_i. \quad (4.28)$$

Если считать начальное значение y_0 произвольной постоянной $y_0 = a_0$, то (4.28) дает общее решение уравнения (4.27).

Пример 11. Найти решение уравнения

$$\Delta y_n = n, \text{ или иначе } y_{n+1} = y_n + n. \quad (4.29)$$

Решение. По формуле (4.28) сразу находим

$$y_n = y_0 + \sum_{i=0}^{n-1} i = y_0 + (1 + 2 + 3 + \dots + (n-1)) = y_0 + \frac{n(n-1)}{2}.$$

Последний член получается после вычисления суммы членов арифметической прогрессии.

Рассмотренное выше уравнение (4.24) – это однородное линейное уравнение первого порядка с переменными коэффициентами. Рассмотрим теперь общий случай неоднородного линейного разностного уравнения первого порядка с переменными коэффициентами

$$\Delta y_n + p_n y_n = q_n. \quad (4.30)$$

Здесь p_n и q_n – известные функции, определенные на множестве $\{0, 1, 2, \dots\}$, а y_n – неизвестная функция, определенная на том же множестве.

Заметим, что способы решения разностных уравнений повторяют аналогичные приемы решения дифференциальных уравнений. По аналогии с дифференциальным уравнением первого порядка с переменными коэффициентами будем искать решение (4.30) в виде произведения двух функций

$$y_n = u_n v_n, \quad (4.31)$$

где функции u_n , v_n пока произвольные и подлежат определению. Из (4.31) найдем разность $\Delta y_n = y_{n+1} - y_n$, входящую в уравнение (4.30):

$$\begin{aligned} \Delta y_n &= y_{n+1} - y_n = u_{n+1} v_{n+1} - u_n v_n = u_{n+1} v_{n+1} - u_{n+1} v_n + u_{n+1} v_n - u_n v_n = \\ &= u_{n+1} (v_{n+1} - v_n) + v_n (u_{n+1} - u_n) = u_{n+1} \Delta v_n + v_n \Delta u_n. \end{aligned}$$

Полученная формула является аналогом формулы для производной от произведения двух функций: $(uv)' = u'v + uv'$. Подставим выражение (4.31) и полученную формулу для Δy_n в уравнение (4.30). Тогда (4.30) примет вид

$$u_{n+1}\Delta v_n + v_n[\Delta u_n + p_n u_n] = q_n. \quad (4.31a)$$

Выберем u_n так, чтобы член в квадратных скобках обратился в нуль

$$\Delta u_n + p_n u_n = 0. \quad (4.32)$$

Тогда, определив u_n из этого уравнения, для функции v_n мы получим простое уравнение, которое следует из (4.31)

$$u_{n+1}\Delta v_n = q_n. \quad (4.33)$$

Решим прежде всего уравнение (4.32). Оно, в сущности, совпадает с решенным ранее уравнением (4.24). Действительно, (4.32) можно записать в виде

$$u_{n+1} - u_n + p_n u_n = u_{n+1} - u_n(1 - p_n) = 0,$$

откуда следует, что

$$u_{n+1} = u_n(1 - p_n). \quad (4.34)$$

Равенство (4.34) совпадает с (4.24), где $\varphi_n = 1 - p_n$. Отсюда следует, что (4.32) [или эквивалентное ему уравнение (4.34)] имеет решение [см. формулу (4.25)]

$$u_n = \prod_{i=0}^{n-1} (1 - p_i). \quad (4.35)$$

Найдя u_n , решим теперь уравнение (4.33), которое можно записать в виде

$$\Delta v_n = \frac{q_n}{u_{n+1}}, \quad (4.36)$$

откуда видно, что это уравнение совпадает с (4.27), которое было решено ранее, причем $\psi_n = q_n/u_{n+1}$. Отсюда по формуле (4.28) можно выписать решение для v_n

$$v_n = v_0 + \sum_{i=0}^{n-1} \left(\frac{q_i}{u_{n+1}} \right). \quad (4.37)$$

Пример 12. Найти общее решение неоднородного уравнения

$$y_{n+1} - \frac{1}{2^n} y_n = 2^{-n^2/2}.$$

Решение. Сравнение данного уравнения с уравнением в общем виде (4.30) показывает, что в нашем случае

$$p_n = 1 - \frac{1}{2^n}, \quad q_n = \frac{1}{2^{n^2/2}}.$$

Беря $y_n = u_n v_n$, определим u_n из уравнения (4.32), что в нашем случае дает уравнение

$$u_{n+1} = \frac{1}{2^n} u_n.$$

Это уравнение было решено в примере 10. Следовательно, решение для u_n имеет вид

$$u_n = C_1 2^{n(n-1)/2}.$$

Функция v_n определяется из уравнения (4.36), которое в данном примере имеет вид

$$\Delta v_n = \frac{2^{-n^2/2}}{C_1 2^{-n(n+1)/2}} = \frac{2^{n/2}}{C_1}.$$

Решение этого уравнения ищем по формуле (4.28)

$$v_n = C_2 + \frac{1}{C_1} \sum_{i=0}^{n-1} (2^{i/2}) = C_2 + \frac{1}{C_1} \cdot \frac{1-2^{n/2}}{1-2^{1/2}}.$$

Следовательно, общим решением исходного неоднородного уравнения будет функция

$$y_n = u_n v_n = \frac{C}{2^{n(n-1)/2}} + \frac{1-2^{n/2}}{2^{n(n-1)/2}(1-\sqrt{2})},$$

где $C = C_1 C_2$.

4.3. Решение линейных разностных уравнений второго порядка с постоянными коэффициентами

Линейным разностным уравнением m -го порядка называется линейное уравнение, связывающее $m+1$ значение сеточной функции.

Общее решение разностного уравнения порядка m – решение, зависящее от m произвольных постоянных, которые можно определить по дополнительным условиям, налагаемым на значения сеточных функций при некоторых значениях аргумента.

Так же как и для дифференциальных уравнений, для линейных разностных уравнений второго порядка существует тип задач, для которых можно получить аналитическое решение. Уравнения этого типа имеют вид

$$a_0 y_{n+2} + a_1 y_{n+1} + a_2 y_n = f_n. \quad (4.38)$$

Здесь a_0, a_1, a_2 – постоянные коэффициенты, причем $a_0 \neq 0$, $a_2 \neq 0$, f_n – заданная сеточная функция, а y_n – искомая функция. Уравнение (4.38) называется неоднородным при $f_n \neq 0$. Если $f_n = 0$, то уравнение называется однородным:

$$a_0 y_{n+2} + a_1 y_{n+1} + a_2 y_n = 0. \quad (4.39)$$

Для разностного уравнения (4.38), как и для аналогичных дифференциальных уравнений, общее решение неоднородного уравнения определяется в виде суммы общего решения однородного уравнения и частного решения неоднородного уравнения. Рассмотрим метод нахождения общего решения однородного уравнения (4.39). Оно представляется в виде

$$Y_n = C_1 y_{1n} + C_2 y_{2n}. \quad (4.40)$$

Здесь Y_n – общее решение однородного уравнения, C_1 и C_2 – произвольные постоянные, а y_{1n} и y_{2n} – фундаментальная система решений уравнения (4.39), которая определяется следующим образом. Решение ищется в виде $y_n = r^n$. Подставив r^n в уравнение (4.39), для определения r получим характеристическое уравнение

$$a_0 r^2 + a_1 r + a_2 = 0, \quad (4.41)$$

корни которого имеют вид

$$r_{1,2} = \frac{-a_1 \pm \sqrt{D}}{2a_0}, \quad D = a_1^2 - 4a_0a_2. \quad (4.42)$$

В зависимости от знака дискриминанта D возникает один из трех возможных случаев:

– если уравнение (4.41) имеет два действительных различных корня r_1 и r_2 , отличные от нуля, то

$$y_{1n} = r_1^n, \quad y_{2n} = r_2^n; \quad (4.43)$$

– если $D = 0$, корни характеристического уравнения действительные и равные $r_1 = r_2 = r \neq 0$, то

$$y_{1n} = r^n, \quad y_{2n} = nr^n; \quad (4.44)$$

– если уравнение имеет комплексно-сопряженные корни

$$r_1 = \alpha + \beta i, \quad r_2 = \alpha - \beta i,$$

где α и β – действительные числа, а $i = \sqrt{-1}$, то в этом случае

$$y_{1n} = \rho^n \cos(\gamma x), \quad y_{2n} = \rho^n \sin(\gamma x),$$

где величины ρ и γ определены следующими соотношениями

$$\rho = \sqrt{\alpha^2 + \beta^2}, \quad \operatorname{tg} \gamma = \beta / \alpha.$$

После построения общего решения однородного уравнения в виде (4.40) для получения общего решения неоднородного уравнения достаточно определить какое-либо частное решение неоднородного уравнения (4.38). Как и в случае дифференциального уравнения, для некоторых специальных видов правой части f_n можно применять метод неопределенных коэффициентов. Более подробно с этими вопросами можно ознакомиться в монографии [5] и работе [10].

5. ОБЫКНОВЕННЫЕ ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ

5.1. Общие понятия

Напомним нужные сведения из курса обыкновенных дифференциальных уравнений. Обыкновенным дифференциальным уравнением называют такое уравнение, которое содержит аргумент x , искомую функцию $y(x)$ и одну или несколько производных от искомой функции:

$$F(x, y, y', \dots, y^{(n)}) = 0. \quad (5.1)$$

Наивысший порядок n входящей в уравнение производной называется порядком дифференциального уравнения. Если удастся в явном виде выразить старшую производную из (5.1), то такое уравнение называется уравнением, разрешенным относительно старшей производной

$$y^{(n)} = \psi(x, y, y', \dots, y^{(n-1)}). \quad (5.2)$$

Решением дифференциального уравнения (5.1) называется такая функция $y(x)$, которая после подстановки ее в уравнение (5.1) превращает его в тождество. Важно отметить, что в общем случае дифференциальное уравнение имеет бесчисленное множество решений. Общее решение уравнения (5.1) содержит n произвольных постоянных и имеет вид

$$y = \varphi(x, C_1, C_2, \dots, C_n). \quad (5.3)$$

Каждое частное решение дифференциального уравнения получается из общего, если произвольные постоянные принимают определенные значения $C_1 = C_{01}, C_2 = C_{02}, \dots, C_n = C_{0n}$. Таким образом, частное решение имеет вид

$$y = \varphi(x, C_{01}, C_{02}, \dots, C_{0n}). \quad (5.4)$$

Для выделения единственного решения дифференциального уравнения (5.1), т.е. для получения частного решения, необходимо задать n произвольных постоянных. При аналитическом решении дифференциальных уравнений методами высшей математики определяется общее решение, зависящее от n произвольных постоянных, после чего они находятся с помощью n дополнительных условий. К сожалению, получить общее решение в аналитическом виде можно лишь в небольшом количестве случаев: для линейных уравнений с постоянными

коэффициентами, для некоторых классов уравнений с переменными коэффициентами (как линейных, так и нелинейных). В большинстве случаев, встречающихся на практике, дифференциальные уравнения приходится решать численно. В этом случае, используя уравнение и дополнительные условия, сразу строится частное решение задачи.

В зависимости от вида дополнительных условий для получения частного решения дифференциального уравнения существуют два различных типа задач: **задача с начальными данными (задача Коши)** и **краевая задача**. В качестве дополнительных условий обычно задаются значения искомой функции и ее производных при некоторых значениях независимой переменной. Если все эти дополнительные условия задаются в одной точке, то такая задача называется задачей Коши, а сами эти условия называются начальными условиями. Если же дополнительные условия задаются более чем в одной точке, то такая задача называется краевой, а сами дополнительные условия называются краевыми (или граничными) условиями. Такое название принято потому, что в большинстве случаев краевые условия задаются в точках, являющихся границами области решения дифференциального уравнения. Приведем примеры формулировки краевых задач и задачи Коши.

Примеры задач Коши:

$$\frac{dy}{dx} = x - \sin^2(yx), \quad x > 0, \quad y(0) = 2; \quad (5.5)$$

$$y'' + 2 \cos xy' - xy = x^2, \quad x > 1, \quad y(1) = 4, \quad y'(1) = 1 \quad (5.6)$$

В первой задаче уравнение имеет первый порядок и поэтому задано одно дополнительное условие. Следующее дифференциальное уравнение имеет второй порядок и поэтому заданы два дополнительных условия в задаче Коши, причем оба в точке $x = 1$.

Примеры краевых задач:

$$y'' + e^x y' - 3y = \cos x, \quad x \in [0, 1], \quad y(0) = 3, y(1) = 2; \quad (5.7)$$

$$y^{(4)} - 2 \sin xy'' + y' = tg^2 x, \quad (5.8)$$

$$x \in [a, b], \quad y(a) = y'(a) = 0, y(b) = y'(b) = 2.$$

Первое уравнение имеет второй порядок и для него заданы два дополнительных условия в граничных точках $x = 0$ и $x = 1$. Второе уравнение имеет четвертый порядок и для него заданы четыре граничных условия в точках $x = a$ и $x = b$.

Среди численных методов решения дифференциальных уравнений чаще всего встречаются **метод конечных разностей** и **метод конечных элементов**. В обоих случаях решение задачи ищется в виде дискретной функции, определенной на дискретном множестве аргументов, и сводится в конечном счете к решению систем линейных алгебраических уравнений (для линейных дифференциальных уравнений) или систем нелинейных уравнений (для нелинейных дифференциальных уравнений). Отличие обоих методов заключается в способе получения систем уравнений. В методе конечных разностей входящие в дифференциальное уравнение производные аппроксимируются конечными разностями, и задача сводится к системе разностных уравнений. В методе конечных элементов краевая задача заменяется некоторой эквивалентной ей задачей, в которой требуется найти минимум функционала для функции, определенной на дискретном множестве точек. Мы будем в основном рассматривать конечно-разностные методы. Рассмотрим сущность этих методов и основные понятия, которые используются при их применении.

5.2. Метод конечных разностей

Сущность метода состоит в следующем. Область непрерывного изменения аргумента (например, отрезок $[a, b]$, на котором определено решение $y(x)$ дифференциального уравнения) заменяется дискретным множеством точек, называемых узлами. Эти узлы образуют разностную сетку. Искомая функция $y(x)$, являющаяся решением дифференциального уравнения и непрерывно зависящая от x , заменяется функцией дискретного аргумента на данной сетке. Эта функция называется **сеточной**. Исходное дифференциальное уравнение заменяется разностным уравнением относительно сеточной функции путем замены входящих в уравнение производных соответствующими конечно-разностными формулами, которые были рассмотрены в разделе 3. Замена дифференциального уравнения разностным называется его **разностной аппроксимацией**. Таким образом, решение дифференциального уравнения сводится к решению разностных уравнений относительно значений сеточной функции в узлах сетки. Для возможности практического использования метода конечных разностей необходимо:

- обосновать возможность замены дифференциального уравнения разностным;
- оценить точность получаемых решений;
- исследовать устойчивость метода.

В конечном счете все перечисленные исследования и оценки

должны помочь сформулировать условия, при выполнении которых решение разностных уравнений сходится к решению дифференциального уравнения с соответствующими граничными или начальными условиями при неограниченном уменьшении шага сетки.

Для компактного описания перечисленных выше вопросов используют операторную запись дифференциальных и разностных уравнений и соответствующих начальных и граничных условий, которая является символическим видом их записи. Приведенные выше в качестве примера дифференциальные уравнения (5.5) – (5.8) можно записать в символическом виде

$$Ly = f(x), \quad x \in G. \quad (5.9)$$

Здесь L – дифференциальный оператор, зависящий от функции $y(x)$ и содержащий функции дифференцирования. Его значение различно для различных дифференциальных уравнений. Функция $f(x)$ называется правой частью операторного уравнения, она зависит от аргумента x . Символом G обозначена область изменения аргумента. В частном случае область G – это отрезок $[a, b]$ или полуось $x > 0$. Дополнительные условия на границе также представляются в операторном виде

$$ly = \varphi(x), \quad x \in Q. \quad (5.10)$$

Здесь l – оператор начальных или граничных условий, который зависит от функции $y(x)$ и, быть может, некоторых ее производных, $\varphi(x)$ – правая часть граничных условий, Q – множество, задающее границу области G (например, $x = 0$ и $x = 1$ или $x = a$ и $x = b$). Поясним операторную запись (5.9), (5.10) на примере краевой задачи (5.7). В этом случае:

$$Ly = y'' + e^x y' - 3y, \quad f(x) = \cos x, \quad G = \{0 \leq x \leq 1\}$$

$$ly = \begin{cases} y(0) \\ y(1) \end{cases} \quad \varphi(x) = \begin{cases} 3 \\ 2 \end{cases} \quad Q = \{x = 0, x = 1\}.$$

При использовании разностного метода дифференциальное уравнение (5.9) заменяется разностным уравнением, которое также можно записать в операторном виде

$$L_h y_h = f_h(x), \quad x \in g_h. \quad (5.11)$$

Здесь L_h – разностный оператор, аппроксимирующий дифференциальный оператор L ; $g_h = \{x_i\}$ ($i = 0, 1, \dots, n$) – дискретное множество аргументов, расположенных на одинаковом расстоянии h друг от друга; y_h – дискретная сеточная функция, определенная на g_h , с помощью которой мы заменяем непрерывную функцию $y(x)$, являющуюся решением дифференциального уравнения (5.11); $f_h(x_i)$ – сеточная функция, определенная на g_h и заменяющая непрерывную функцию $f(x)$. Индекс h в (5.11) подчеркивает, что величина шага сетки является параметром разностной задачи. Следовательно, (5.11) можно рассматривать как целое семейство разностных уравнений, которые зависят от параметра h . При рассмотрении формул численного дифференцирования в главе 3 отмечалось, что погрешность аппроксимации производной в некоторой точке x_i может быть представлена в виде $O(h^k)$. В этом случае, имея в виду весь дифференциальный оператор (5.11), говорят, что в точке x_i имеет место **аппроксимация k -го порядка**.

При численном решении дифференциальных уравнений требуется оценить погрешность аппроксимации не в одной точке, а на всей сетке g_h , т.е. в узлах $x_0, x_1, \dots, x_n$. В качестве погрешности аппроксимации ε_h на всей сетке можно принять некоторую интегральную величину, связанную с погрешностями аппроксимации в отдельных узлах. Наиболее часто в качестве такой величины принимается одна из следующих величин:

$$\varepsilon_h = \max_i |\varepsilon(x_i)|; \quad (5.12)$$

$$\varepsilon_h = \left[\sum_i \varepsilon^2(x_i) \right]^{1/2}. \quad (5.13)$$

Величина (5.12) означает, что в качестве меры интегральной аппроксимации на всем отрезке принимается наибольшая из погрешностей в отдельных узлах сетки, а (5.13) означает выбор среднеквадратичной величины.

Наряду с аппроксимацией (5.11) дифференциального уравнения

(5.9) вводится аппроксимация начальных или граничных условий (5.10), которые можно записать в виде

$$l_h y_h = \varphi_h(x), \quad x \in q_h. \quad (5.14)$$

Здесь q_h – граничные узлы сетки, а индекс h означает зависимость разностной аппроксимации граничных условий от значения шага h .

Совокупность разностных уравнений (5.11) и (5.14), аппроксимирующих исходное дифференциальное уравнение и дополнительные условия на границах, называется **разностной схемой**.

Главный вопрос, на который необходимо ответить при использовании разностных схем, заключается в том, будет ли решение разностной задачи y_h сходиться к точному решению дифференциальной задачи $y(x)$ во всех узловых точках сетки при неограниченном измельчении сетки, т.е. при $h \rightarrow 0$. Для ответа на этот вопрос необходимо рассмотреть такие понятия, как сходимость, устойчивость и аппроксимация разностной схемы.

Выше была введена интегральная величина ε_h погрешности аппроксимации дифференциального уравнения разностным на всем отрезке [формулы (5.13) или (5.14)]. Аналогично можно ввести величину погрешности замены краевых или начальных условий их разностными соотношениями. Далее под величиной ε_h мы будем понимать интегральную величину общей погрешности, вносимой как разностной схемой, где учитываются ошибки, вносимые заменой дифференциального уравнения разностным, так и заменой дополнительных условий (начальных или граничных) конечно-разностными выражениями.

Будем говорить, что *разностная схема (5.11), (5.14) аппроксимирует дифференциальное уравнение (5.9) и дополнительные условия на границах (5.10), если*

$$\varepsilon_h \rightarrow 0 \text{ при } h \rightarrow 0. \quad (5.15)$$

Если при этом

$$|\varepsilon_h| \leq ch^k, \quad (5.16)$$

где $c > 0$ и $k > 0$ – некоторые постоянные, не зависящие от величины шага h , то говорят, что *аппроксимация имеет порядок k относительно шага сетки h* .

Разностная схема называется **устойчивой**, если малому изменению входных данных соответствует малое изменение решения разност-

ных уравнений. Входными данными являются коэффициенты в разностных уравнениях, их правые части и начальные или граничные условия. В противном случае разностная схема называется **неустойчивой**.

Будем говорить, что решение разностной краевой задачи (5.11), (5.14) сходится к решению дифференциальной задачи (5.9), (5.10), если

$$\max_i |y(x_i) - y_h(x_i)| \rightarrow 0 \quad \text{при } h \rightarrow 0. \quad (5.17)$$

Здесь $y(x_i)$ – точное решение дифференциальной задачи в точке $x = x_i$, а $y_h(x_i)$ – решение разностной задачи в этой же точке.

Основной результат, полученный в теории разностных схем, заключается в том, что *если разностная схема аппроксимирует исходную дифференциальную задачу и является устойчивой, то решение, полученное по разностной схеме, сходится к решению дифференциальной задачи* при неограниченном измельчении сетки, т.е. при $h \rightarrow 0$. Другими словами, из устойчивости и аппроксимации разностной схемы следует ее сходимости. Это означает, что трудную задачу исследования сходимости разностной схемы к решению дифференциальной задачи можно заменить более простыми задачами изучения аппроксимации схемы и исследования ее устойчивости.

Замечание 1. Можно показать, что если устойчивая разностная схема имеет порядок аппроксимации k , т.е. выполняется условие (5.16), то

$$\max_i |y(x_i) - y_h(x_i)| \leq Mh^k, \quad (5.18)$$

где M – константа, не зависящая от h . Это фактически означает, что порядок сходимости решения разностной задачи к решению дифференциальной задачи в зависимости от h будет таким же, как порядок аппроксимации.

Замечание 2. В приведенных выше результатах в качестве показателя близости двух функций (решений дифференциальной и разностной задачи) рассматривалась интегральная характеристика типа (5.12). Аналогичные результаты могут быть получены и в случае меры близости решений, определяемых соотношением (5.13), которое характеризует приближение “в среднем”.

5.3. Решение задачи Коши

5.3.1. Основные понятия

Начнем рассмотрение с дифференциального уравнения первого порядка, разрешенного относительно производной

$$\frac{dy}{dx} = f(x, y), \quad (5.19)$$

удовлетворяющего начальному условию (задача Коши)

$$y(x_0) = y_0. \quad (5.20)$$

Решением задачи Коши является функция $y(x)$, которая удовлетворяет уравнению (5.19) и начальному условию (5.20). Геометрически это означает, что из множества функций, удовлетворяющих уравнению (5.19), мы находим такое частное решение, которое проходит через точку с координатами (x_0, y_0) (рис. 5.1). Используем для решения задачи (5.19), (5.20) конечно-разностный метод. Введем узловые точки $x_0, x_1, \dots, x_n$, расположенные с постоянным шагом h : $x_{i+1} - x_i = x_i - x_{i-1} = \text{const} = h$. Рассмотрим сеточную функцию $y_0, y_1, \dots, y_n$, определенную в узлах сетки. Используя различные разностные аппроксимации для производной dy/dx , мы можем получить различные разностные уравнения. Их общий вид

$$y_{i+1} = \varphi(x_i, h, y_{i+1}, y_i, y_{i-1}, \dots, y_{i-k+1}). \quad (5.21)$$

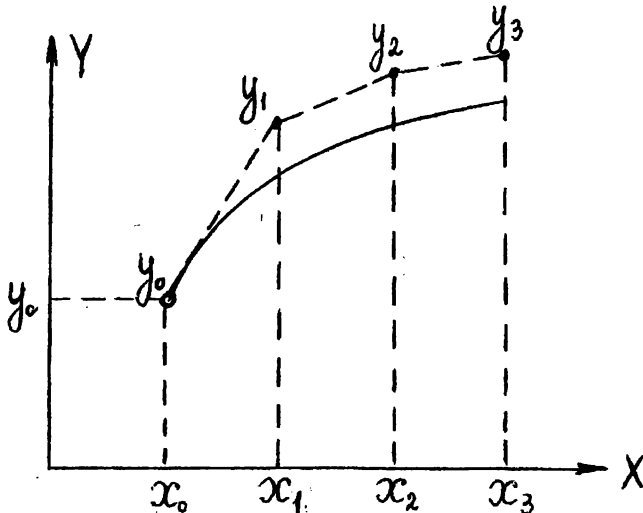


Рис 5.1

Конкретный вид и тип разностной схемы зависят от того, какие из выписанных аргументов в правой части уравнения в ней в действительности присутствуют. Если в правой части (5.21) отсутствует величина y_{i+1} , то схема называется явной. Это означает, что значение y_{i+1} , стоящее в левой части, явно вычисляется по предыдущим значениям $y_i, y_{i-1}, \dots, y_{i-k+1}$. В случае, когда $k = 1$, мы получаем **одношаговый метод**

$$y_{i+1} = \varphi(x_i, h, y_i). \quad (5.22)$$

В одношаговом методе решение в очередной точке вычисляется с использованием значения функции в предыдущей точке.

При $k = 2$ получаем двухшаговый метод

$$y_{i+1} = \varphi(x_i, h, y_i, y_{i-1}). \quad (5.23)$$

В одношаговом методе значение начального условия $y(x_0) = y_0$ позволяет последовательно определить значения сеточной функции во всех последовательных узлах по формуле (5.22): $y_1 = \varphi(x_0, h, y_0)$, $y_2 = \varphi(x_1, h, y_1)$, $y_3 = \varphi(x_2, h, y_2)$, ..., $y_n = \varphi(x_{n-1}, h, y_{n-1})$. В двухшаговом методе (5.23) для вычисления значения y_{i+1} требуется знание двух предыдущих значений функции y_i, y_{i-1} , и поэтому одного начального значения y_0 для определения следующего значения y_1 здесь уже недостаточно. Аналогично обстоит дело и для k -шаговых методов при $k = 3$.

В **многошаговом методе решения задачи Коши** для вычисления очередного решения используется несколько предыдущих уже полученных значений решения в нескольких точках.

Если в правую часть (5.21) входит искомое значение y_{i+1} , то метод называется неявным. В этом случае (5.21) не позволяет явно вычислить величину y_{i+1} и для ее нахождения приходится использовать итерационные методы.

5.3.2. Метод Эйлера

Рассмотрим наиболее простой явный одношаговый метод, который называют **методом Эйлера**. Будем решать уравнение (5.19) с начальным условием (5.20)

$$y' = f(x, y), \quad y(x_0) = y_0$$

на введенной ранее сетке с постоянным шагом h . Запишем в начальной точке x_0 значение первой производной с помощью правой разности

$$y'(x_0) = \frac{y_1 - y_0}{h} + O(h).$$

Используя это выражение, запишем разностное уравнение, приближенно аппроксимирующее уравнение (5.19) в точке x_0

$$\frac{y_1 - y_0}{h} + O(h) = f(x_0, y_0).$$

Выражая из этого уравнения неизвестную величину y_1 , получим

$$y_1 = y_0 + hf(x_0, y_0) + O(h^2). \quad (5.24)$$

Поскольку x_0 и y_0 известны из начальных условий, то, пренебрегая последним членом $O(h^2)$, который определяет погрешность аппроксимации дифференциального уравнения разностным, получим формулу метода Эйлера

$$y_1 = y_0 + hf(x_0, y_0). \quad (5.25)$$

Так как $x_1 = x_0 + h$, то (5.25) определяет следующую точку (x_1, y_1) , которую мы можем принять за начальную для следующего шага

$$y_2 = y_1 + hf(x_1, y_1).$$

Аналогично может быть найдено значение сеточной функции в любом узле сетки

$$y_{i+1} = y_i + hf(x_i, y_i). \quad (5.26)$$

Таким образом, начиная с начальной точки (x_0, y_0) , мы можем последовательно определить значения сеточной функции y_i в любых узлах x_i , используя явную одношаговую формулу Эйлера (5.26). Как следует из (5.24), метод Эйлера имеет на каждом шаге локальную ошибку $O(h^2)$. Для определения величины y_n нам придется сделать n шагов

по формуле (5.26), на каждом из которых погрешность имеет порядок $O(h^2)$, и, таким образом, за n шагов погрешность будет $\approx nO(h^2) \cong \frac{1}{h}O(h^2) \approx O(h)$. Следовательно, глобальная погрешность за

$\bar{n}$ шагов имеет погрешность первого порядка $O(h)$. Согласно формуле для погрешности, для ее уменьшения надо уменьшать шаг h , однако в этом случае увеличивается общее количество шагов и могут накапливаться погрешности округления.

Метод Эйлера имеет простую геометрическую интерпретацию. Так как $f(x_0, y_0) = y'_0$, то из геометрического смысла производной следует, что $f(x_0, y_0) = \operatorname{tg} \alpha$, где α – угол наклона касательной к кривой $y(x)$, являющейся решением дифференциального уравнения в точке (x_0, y_0) . Поэтому результат первого шага y_1 в методе Эйлера [формула(5.25)] есть пересечение касательной к решению дифференциального уравнения в точке x_0 с прямой $x = x_1$ (рис.5.1). Следовательно, метод Эйлера заменяет кривую $y(x)$ на отрезке $[x_0, x_1]$ прямой линией, являющейся касательной к функции $y(x)$ в точке (x_0, y_0) , а истинное решение $y(x)$ на отрезке $[x_0, x_n]$ ломаной (рис.5.1), которая с ростом x может все дальше уходить от решения $y(x)$. Отсюда же видно, как можно усовершенствовать метод Эйлера. Для этого надо учесть изменение угла наклона касательной к кривой $y(x)$ при перемещении от точки x_0 к точке x_1 , а не брать этот угол постоянным, равным $\operatorname{arctg}(f(x_0, y_0))$, как это сделано в методе Эйлера.

5.3.3. Метод Эйлера с пересчетом

Одним из возможных улучшений метода Эйлера является **метод Эйлера с пересчетом**, который для получения величины y_1 использует два этапа. На первом этапе определяется предварительное значение $\bar{y}_1$ по методу Эйлера

$$\bar{y}_1 = y_0 + hf(x_0, y_0). \quad (5.27)$$

Имея теперь точки (x_0, y_0) и $(x_1, \bar{y}_1)$, мы определим значения

производной $y'(x) = f(x, y)$ в этих точках, которые будут равны $f(x_0, y_0)$ и $f(x_1, \bar{y}_1)$, и возьмем среднее значение этих производных $f_c = \frac{1}{2}[f(x_0, y_0) + f(x_1, \bar{y}_1)]$. На втором этапе определяется уточненное значение y_1 по формуле

$$y_1 = y_0 + hf_c = y_0 + h/2[f(x_0, y_0) + f(x_1, \bar{y}_1)]. \quad (5.28)$$

Геометрически формула (5.28) означает, что мы опять получаем решение на отрезке $[x_0, x_1]$ в виде прямой, но теперь наклон этой прямой определяется средним значением производных в начальной точке x_0 и конечной точке x_1 , а не только в начальной точке, как это сделано в методе Эйлера. В общем виде один шаг метода Эйлера с пересчетом можно записать следующим образом:

$$\bar{y}_{i+1} = y_i + hf(x_i, y_i); \quad (5.29)$$

$$y_{i+1} = y_i + \frac{h}{2}[f(x_i, y_i) + f(x_{i+1}, \bar{y}_{i+1})]. \quad (5.30)$$

Метод Эйлера с пересчетом является частным случаем общего подхода к решению задачи Коши, называемого методом прогноза и коррекции. В данном случае определение предварительного значения $\bar{y}_{i+1}$ является прогнозом истинной величины $y_{i+1} = y(x_{i+1})$, где $y(x)$ – точное решение задачи. Второй этап метода является корректировкой предварительного прогноза $\bar{y}_{i+1}$ и дает определение более точного значения y_{i+1} , которое гораздо ближе к точному решению. Можно показать, что локальная погрешность метода Эйлера с пересчетом будет иметь порядок $O(h^3)$, а глобальная погрешность на отрезке $[x_0, x_n]$ – порядок $O(h^2)$. Этот метод является первой итерацией более общего подхода, определяемого формулой

$$y_{i+1} = y_i + \frac{h}{2}[f(x_i, y_i) + f(x_{i+1}, y_{i+1})]. \quad (5.31)$$

Метод (5.31) является неявным, так как искомая величина y_{i+1} входит как в левую, так и в правую часть уравнения. Можно построить итерационный метод определения y_{i+1} из (5.31), но поскольку такая

итерационная процедура должна применяться на каждом шаге решения задачи Коши, то метод будет достаточно громоздким.

Метод Эйлера с пересчетом реализует один шаг метода итераций. С помощью этого метода можно проводить контроль точности решения путем сравнения значений $\bar{y}_{i+1}$ и y_{i+1} и выбора на основании этого величины шага h в каждом узле x_i . Для этого необходимо на каждом шаге вычислять величину $|y_{i+1} - \bar{y}_{i+1}|$. Если эта величина сравнима с погрешностями вычислений, то шаг h нужно увеличить; в противном случае, если эта разность слишком велика (например, $|y_{i+1} - \bar{y}_{i+1}| > 0,01y_{i+1}$), значение h следует уменьшить. С помощью процедур такого типа можно реализовать автоматический выбор шага при решении задачи Коши.

5.3.4. Метод Рунге-Кутты

Помимо рассмотренных существуют и другие явные одношаговые методы. На практике чаще всего используется метод Рунге-Кутты, который для гладких функций обладает очень высокой локальной точностью $O(h^5)$ и суммарной точностью на отрезке $[x_0, x_n]$ порядка $O(h^4)$. Для того чтобы привести формулы этого метода, заметим, что метод Эйлера с пересчетом можно записать в следующей компактной форме

$$y_{i+1} = y_i + \Delta y_i. \quad (5.32)$$

Тогда для метода Эйлера

$$\Delta y_i = h_i f(x_i, y_i), \quad (5.33)$$

а для метода Эйлера с пересчетом

$$\Delta y_i = \frac{h}{2}(K_1 + K_2); \quad (5.34)$$

$$K_1 = f(x_i, y_i), \quad K_2 = f(x_i + h_i, y_i + K_1 h_i). \quad (5.35)$$

В формулах (5.35) производные K_1 и K_2 вычисляются в начальной и конечной точках интервала $[x_i, x_i + h_i]$. Метод Рунге-Кутты является обобщением такого подхода. В нем производные вычисляются как в крайних точках, так и середине интервала $[x_i, x_{i+1}]$, а затем берется

некоторое среднее значение

$$\Delta y_i = \frac{h}{6}(K_1 + 2K_2 + 2K_3 + K_4); \quad (5.36)$$

$$\begin{cases} K_1 = f(x_i, y_i), & K_2 = f\left(x_i + \frac{h_i}{2}, y_i + \frac{K_1 h_i}{2}\right); \\ K_3 = f\left(x_i + \frac{h_i}{2}, y_i + \frac{K_2 h_i}{2}\right), & K_4 = f(x_i + h_i, y_i + K_3 h_i). \end{cases} \quad (5.37)$$

Коэффициенты при $K_i (i = 1, 2, 3, 4)$ в (5.36) определены теоретически из условия получения наиболее высокой точности метода. Метод Рунге-Кутты требует четырехкратного вычисления производных, т.е. функции $f(x, y)$, по формулам (5.37). Однако это окупается его весьма высокой точностью, что позволяет решать дифференциальное уравнение с гораздо большим шагом, чем в методах Эйлера и Эйлера с пересчетом, при сравнимой точности решений. Кроме того, метод Рунге-Кутты легко позволяет в процессе решения изменять шаг h_i в зависимости от получаемой гладкости решения: если получаемое решение сильно растет, то шаг h_i уменьшается, если же решение изменяется слабо при переходе от x_i к $x_{i+1} = x_i + h_i$, то шаг h_i будет увеличиваться. В различных системах математического обеспечения ЭВМ имеются хорошие и эффективные программы, реализующие метод Рунге-Кутты с переменным шагом и заданной точностью вычислений.

Приведем пример решения задачи Коши с помощью трех рассмотренных методов и сравним точность получаемых результатов.

Пример 13. Решить задачу Коши

$$y' = 2(x^2 + y); \quad y(0) = 1; \quad x \in [0, 1]; \quad h = 0,1. \quad (5.38)$$

Решение. Прежде всего заметим, что эту задачу можно решить известными методами из курса высшей математики. Ее точное решение с учетом начального условия имеет вид $y = 1,5e^{2x} - x^2 - x - 0,5$. Для проведения численных расчетов заметим, что в нашем случае $f(x, y) = 2(x^2 + y)$.

а) Метод Эйлера [формула(5.26)]. Определяем результат первого шага

$$y_1 = y_0 + hf(x_0, y_0) = 1 + 0,1 \cdot [2 \cdot (0 + 1)] = 1,2.$$

На втором шаге получаем

$$y_2 = y_1 + hf(x_1, y_1) = 1,2 + 0,1 \cdot [2 \cdot (0,1^2 + 1,2)] = 1,4420.$$

Результаты счета при других x приведены в таблице 5.1.

б) Метод Эйлера с пересчетом [формулы (5.29), (5.30)]. Результаты первого шага:

$$\bar{y}_1 = y_0 + hf(x_0, y_0) = 1 + 0,1 \cdot [2 \cdot (0 + 1)] = 1,2;$$

$$y_1 = y_0 + \frac{h}{2} [f(x_0, y_0) + f(x_1, \bar{y}_1)] = 1 + \frac{0,1}{2} [2 \cdot (0 + 1) + 2 \cdot (0,1^2 + 1,2)] = 1,2210.$$

Результаты второго шага:

$$\bar{y}_2 = y_1 + hf(x_1, y_1) = 1,221 + 0,1 \cdot [2 \cdot (0,1^2 + 1,2210)] = 1,4672;$$

$$\begin{aligned} y_2 &= y_1 + \frac{h}{2} [f(x_1, y_1) + f(x_2, \bar{y}_2)] = \\ &= 1,221 + \frac{0,1}{2} [2 \cdot (0,1^2 + 1,221) + 2 \cdot (0,2^2 + 1,4672)] = 1,4923. \end{aligned}$$

Результаты счета при других x приведены в табл. 5.1.

в) Метод Рунге-Кутты [формулы (5.36), (5.37)]. Приведем результаты вычисления одного шага:

$$K_1 = f(x_0, y_0) = 2 \cdot [0 + 1] = 2;$$

$$K_2 = f\left(x_0 + \frac{h}{2}, y_0 + \frac{K_1 h}{2}\right) = 2 \cdot \left[\left(0 + \frac{0,1}{2}\right)^2 + 1 + \frac{2 \cdot 0,1}{2}\right] = 2,205;$$

$$K_3 = f\left(x_0 + \frac{h}{2}, y_0 + \frac{K_2 h}{2}\right) = 2 \cdot \left[\left(0 + \frac{0,1}{2}\right)^2 + 1 + \frac{2,205 \cdot 0,1}{2}\right] = 2,2255;$$

$$K_4 = f(x_0 + h, y_0 + K_3 h) = 2 \cdot [(0,1)^2 + 1 + 2,2255 \cdot 0,1] = 2,4651;$$

$$y_1 = y_0 + \frac{h}{6}(K_1 + 2K_2 + 2K_3 + K_4) =$$

$$= 1 + \frac{0,1}{6}(2 + 2 \cdot 2,205 + 2 \cdot 2,2255 + 2,4651) = 1,2221.$$

В табл. 5.1 приведены результаты всех трех методов на интервале $[0, 1]$ с постоянным шагом $h = 0,1$ и значения точного решения.

Таблица 5.1

x_i	Метод Эйлера	Метод Эйлера с пересчетом	Метод Рунге-Кутта	Точное решение
0,1	1,2000	1,2210	1,2221	1,2221
0,2	1,4420	1,4923	1,4977	1,4977
0,3	1,7384	1,8284	1,8432	1,8432
0,4	2,1041	2,2466	2,2783	2,2783
0,5	2,5569	2,7680	2,8274	2,8274
0,6	3,1183	3,4176	3,5201	3,5202
0,7	3,8139	4,2257	4,3927	4,3928
0,8	4,6747	5,2288	5,4894	5,4895
0,9	5,7376	6,4004	6,8643	6,8645
1,0	7,0472	8,0032	8,5834	8,5836

Результаты, приведенные в таблице, показывают очень высокую точность метода Рунге-Кутта и нарастание погрешностей метода Эйлера с увеличением x_i . При $x = 1$ относительные погрешности $|y - y_*|/y_*$,

где y_* – точное решение, равны: для метода Эйлера $\approx 0,18$; для метода Эйлера с пересчетом $\approx 0,068$; для метода Рунге-Кутта $\approx 0,000023$.

Сравнивая работу, необходимую для вычисления одного шага по каждому из методов, отметим, что для метода Эйлера необходимо один раз вычислить правую часть уравнения $f(x, y)$, для метода с пересчетом – два раза, а для метода Рунге-Кутта – четыре раза. Однако лишняя работа в методе Рунге-Кутта с лихвой окупается его высокой точностью, что позволяет проводить вычисления с существенно большими шагами, чем в методе Эйлера или в методе Эйлера с пересчетом.

5.3.5. Многошаговые методы

Недостатком метода Рунге-Кутты является необходимость вычислять четыре раза правую часть уравнения $f(x, y)$. От этого недостатка свободны **многошаговые методы**, которые для вычисления значения решения y_{i+1} используют значения решения не в одной точке y_i , как это делается в одношаговых методах, а в нескольких предыдущих точках $y_i, y_{i-1}, y_{i-2}, \dots, y_{i-k+1}$. Идея построения многошаговых методов заключается в следующем. Рассмотрим исходное уравнение $y' = f(x, y)$ и проинтегрируем обе части этого уравнения по x на отрезке $[x_i, x_{i+1}]$. Интеграл по левой части вычисляется сразу

$$\int_{x_i}^{x_{i+1}} y' dx = y_{i+1} - y_i,$$

а интеграл по правой части – путем предварительного построения интерполяционного многочлена $P_{k-1}(x)$ степени $k - 1$, который аппроксимирует функцию $f(x, y)$ на отрезке $[x_i, x_{i+1}]$ по значениям этой функции в нескольких предыдущих точках, решения в которых уже получены: $f(x_{i-1}, y_{i-1}), \dots, f(x_{i-k+1}, y_{i-k+1})$. Тогда

$$y_{i+1} - y_i = \int_{x_i}^{x_{i+1}} f(x, y) dx \approx \int_{x_i}^{x_{i+1}} P_{k-1}(x) dx.$$

На основе этой формулы можно строить различные многошаговые методы любого порядка точности. Порядок точности зависит от степени интерполяционного многочлена, для построения которого используются значения сеточной функции $y_i, y_{i-1}, \dots, y_{i-k+1}$, вычисленные на k предыдущих шагах. Приведем без вывода две расчетные формулы из семейства явных многошаговых методов Адамса-Башфорта.

Двухшаговый метод второго порядка точности записывается в виде

$$y_{i+1} = y_i + \frac{h}{2} [3f(x_i, y_i) - f(x_{i-1}, y_{i-1})]. \quad (5.39)$$

Более точный четырехшаговый метод четвертого порядка точности описывается формулой

$$y_{i+1} = y_i + \frac{h}{24} [55f(x_i, y_i) - 59f(x_{i-1}, y_{i-1}) + \\ + 37f(x_{i-2}, y_{i-2}) - 9f(x_{i-3}, y_{i-3})]. \quad (5.40)$$

Для вычисления y_{i+1} по формуле (5.39) необходимо знать решения в двух предыдущих точках y_i и y_{i-1} , а по (5.40) – решения в четырех предыдущих точках $y_i, y_{i-1}, y_{i-2}, y_{i-3}$. После этого можно вычислить правые части уравнения $f(x, y)$ в этих точках $f(x_i, y_i), f(x_{i-1}, y_{i-1}), f(x_{i-2}, y_{i-2})$ и воспользоваться формулами (5.39) и (5.40) для определения y_{i+1} . Отметим, что переходя к вычислению на следующем шаге величины y_{i+2} , достаточно вычислить функцию $f(x, y)$ только в одной точке $f(x_{i+1}, y_{i+1})$, так как предыдущие значения $f(x_i, y_i), f(x_{i-1}, y_{i-1}), f(x_{i-2}, y_{i-2})$ уже вычислены.

Недостаток многошаговых схем заключается в том, что невозможно начать счет по одному лишь известному значению y_0 , которое задано в начальном условии. Например, в (5.40) расчет можно начинать лишь со значения y_4 в точке x_4 . Необходимые для такого расчета значения y_0, y_1, y_2 и y_3 должны быть вычислены каким-либо другим методом (например, методом Эйлера или Рунге-Кутты). При этом начальные точки надо вычислить очень точно, так как иначе полученные в начальных точках погрешности будут влиять на решение задачи во всех остальных точках. Другой недостаток многошаговых методов заключается в том, что они, как правило, не позволяют изменять шаг в процессе расчета (без существенного усложнения алгоритма расчета), чтобы учесть характер получаемого решения, подстраивая шаг h к скорости изменения решения.

5.3.6. Сравнительная оценка методов и метод Рунге повышения точности результатов

Большое количество методов решения задачи Коши отражает факт большого разнообразия таких задач, встречающихся на практике. Рассмотрим основные критерии, которыми надо руководствоваться при выборе метода для решения конкретной задачи. Если решение дифференциального уравнения имеет высокую гладкость (для решения

существует достаточное количество производных во всех точках), то целесообразно использовать методы, имеющие высокую точность, т.е. одношаговые методы типа Рунге-Кутты или многошаговые методы типа Адамса-Башфорта.

Выбор между этими двумя типами методов определяется двумя обстоятельствами: сложностью вычисления правой части уравнения $f(x, y)$; разнообразием поведения решения дифференциального уравнения на отрезке, где ищется решение. Если правая часть $f(x, y)$ имеет сложный вид и требует большого количества арифметических операций, то целесообразно использовать методы, где количество вычислений функции $f(x, y)$ будет как можно меньше. В методе Рунге-Кутты для получения y_{i+1} требуется четыре раза вычислять функцию $f(x, y)$ при различных значениях ее аргументов, а в методе Адамса-Башфорта – лишь один раз. В этом случае имеет смысл с помощью метода Рунге-Кутты вычислить с высокой точностью первые шаги решения y_0, y_1, y_2, y_3 , а затем перейти на многошаговый метод Адамса-Башфорта. Второе обстоятельство связано с поведением решения на интервале нахождения решения. Если функция на одних участках сильно возрастает или убывает, а на других участках изменяется незначительно, то требуется применять адаптивные алгоритмы, в которых шаг h сеточной области изменяется в зависимости от поведения решения: при сильном изменении решения величина h уменьшается, а при слабом изменении – увеличивается. Для этих случаев гораздо лучше приспособлен метод Рунге-Кутты, для которого уменьшение или увеличение шага реализуется очень просто. Для метода Адамса-Башфорта уменьшение или увеличение шага h требует серьезной перестройки метода расчета.

Особо следует сказать о задачах, в которых функция $f(x, y)$ и решение являются негладкими, а иногда и разрывными функциями. Особенно часто такая ситуация возникает в задачах регулирования и управления, когда различные воздействия могут иметь прерывистый характер. Выше отмечалось, что применение методов высокой точности накладывает серьезные ограничения на гладкость получаемого решения задачи: методы высокой точности могут быть реализованы только для достаточно гладких решений. Отсюда следует, что для задач с негладкими решениями приходится применять простые методы типа метода Эйлера, которые обладают невысокой точностью.

Очень важным является вопрос о том, как можно повысить точность получаемых разностными методами результатов, не прибегая к чрезмерному уменьшению шага сетки h , так как это увеличивает требуемое машинное время и приводит к нарастанию погрешностей

округления. Одним из эффективных способов повышения точности получения решения без существенного увеличения машинного времени является применение метода Рунге. Он заключается в том, что проводятся две серии расчетов одной и той же задачи с шагами h и $h/2$. На основании этих результатов y_h и $y_{h/2}$ проводится приближенная оценка погрешности решения и строится уточненное решение. Пусть разностная схема аппроксимирует дифференциальную задачу с погрешностью p . Это означает, что

$$y(x) - y_h(x) = C(x)h^p + O(h^{p+1}). \quad (5.41)$$

Здесь $y(x)$ – точное решение дифференциальной задачи, $y_h(x)$ – решение разностного уравнения, $C(x)$ – неизвестная функция, которая зависит от некоторых констант, значений производной решения в некоторой неизвестной точке, и не зависит от шага h , $O(h^{p+1})$ – величина более высокого порядка малости по сравнению с h^p . Оценим величину погрешности решения $R = C(x)h^p$ с помощью двух расчетов с шагами h и $h/2$. При расчете с шагом h имеет место соотношение (5.41). При расчете с шагом $h/2$ (5.41) примет вид

$$y(x) - y_{h/2}(x) = C(x)\left(\frac{h}{2}\right)^p + O\left[\left(\frac{h}{2}\right)^{p+1}\right] = C(x)\left(\frac{1}{2}\right)^p h^p + O(h^{p+1}). \quad (5.42)$$

Вычтем из (5.42) почленно соотношение (5.41) и учтем, что

$$O\left[\left(\frac{h}{2}\right)^{p+1}\right] \approx O(h^{p+1}).$$

Получим, пренебрегая малой величиной $O(h^{p+1})$,

$$-y_h(x) + y_{h/2}(x) = C(x)\left[\frac{h^p}{2^p} - h^p\right] = C(x)h^p\left[\frac{-1 + 2^p}{2^p}\right].$$

Отсюда можно определить выражение для главного члена погрешности $R = C(x)h^p$

$$R = C(x)h^p = \frac{2^p(y_{h/2} - y_h)}{2^p - 1}. \quad (5.43)$$

Подставив это выражение в формулу (5.42), получим уточненное значение полученного ранее решения $y_{h/2}(x)$

$$\bar{y}(x) = y_{h/2} + \frac{y_{h/2} - y_h}{2^p - 1} + O(h^{p+1}). \quad (5.44)$$

Таким образом, решение задачи на двух сетках позволяет на порядок повысить точность результатов. Приведем пример использования метода Рунге.

Пример 14. Для задачи Коши

$$y' = 2(x^2 + y); \quad y(0) = 1; \quad x \in [0, 1]$$

определить уточненное решение в точке $x = 0,2$ методом Эйлера, получив решения с шагом $h = 0,1$ и $h = 0,2$. По формуле метода Эйлера $y_{i+1} = y_i + hf(x_i, y_i)$ получим для шага $h = 0,2$

$$y_1 = y_0 + hf(x_0, y_0) = 1 + 0,2 \cdot [2 \cdot (0 + 1)] = 1,4.$$

Таким образом, $y_h(0,2) = 1,4$. Сделаем теперь два шага с вдвое меньшим шагом $h = 0,1$:

$$y_1 = y_0 + hf(x_0, y_0) = 1 + 0,1 \cdot [2 \cdot (0 + 1)] = 1,2;$$

$$y_2 = y_1 + hf(x_1, y_1) = 1,2 + 0,1 \cdot [2 \cdot (0,1 + 1,2)] = 1,442.$$

Таким образом, $y_{h/2} = 1,442$.

Так как метод Эйлера имеет первый порядок точности, то для него $p = 1$ и из формулы (5.44) получаем

$$\bar{y}(x) = y_{h/2} + y_{h/2} - y_h + O(h^2) = 2y_{h/2} - y_h + O(h^2).$$

Подставив полученные значения $y_h(0,2) = 1,4$ и $y_{h/2} = 1,442$, получим уточненное значение $\bar{y}(0,2)$

$$\begin{cases} y_{1,i+1} = y_{1,i} + hf_1(x_i, y_{1,i}, y_{2,i}, \dots, y_{n,i}) \\ y_{2,i+1} = y_{2,i} + hf_2(x_i, y_{1,i}, y_{2,i}, \dots, y_{n,i}) \\ \vdots \\ y_{n,i+1} = y_{n,i} + hf_n(x_i, y_{1,i}, y_{2,i}, \dots, y_{n,i}) \end{cases} \quad (5.47)$$

Здесь $y_{1,i} = y_1(x_i)$, $y_{1,i+1} = y_1(x_{i+1})$, ..., $y_{n,i} = y_n(x_i)$, $y_{n,i+1} = y_n(x_{i+1})$. В правой части (5.47) стоят значения функций в исходной точке x_i , а слева новые вычисленные значения функций в следующей точке x_{i+1} . Используя начальные условия (5.46), можно последовательно определить значения функций y_k ($k=1,2,...,n$) во всех последующих узлах сетки x_i ($i=1,2,...$) по формулам (5.47). Формулы, аналогичные (5.47), можно записать и для метода Эйлера с пересчетом и для метода Рунге-Кутты. Для знакомства с ними можно обратиться к работам [1,3].

Рассмотрим теперь дифференциальное уравнение n -го порядка, записанное в виде, разрешенном относительно старшей производной:

$$y^{(n)} = f(x, y, y', \dots, y^{(n-1)}). \quad (5.48)$$

Верхний индекс в (5.48) указывает порядок производной. Начальные условия в задаче Коши запишем в виде

$$y(x_0) = z_0, y'(x_0) = z_1, y''(x_0) = z_2, \dots, y^{(n-1)}(x_0) = z_{n-1}. \quad (5.49)$$

Покажем, что уравнение (5.48) можно свести к системе n уравнений первого порядка. Введем следующие обозначения:

$$y = y_1, y' = y_2, y'' = y_3, y''' = y_4, \dots, y^{(n-1)} = y_n. \quad (5.50)$$

Используя эти обозначения, уравнение (5.48) можно записать в виде системы уравнений

$$\begin{cases} y'_1 = y_2 \\ y'_2 = y_3 \\ \dots\dots\dots \\ y'_{n-1} = y_n \\ y'_n = f(x, y_1, y_2, \dots, y_n) \end{cases} \quad (5.51)$$

Начальные условия (5.50) запишутся в виде

$$y_1(x_0) = z_0, y_2(x_0) = z_1, y_3(x_0) = z_2, \dots, y_n(x_0) = z_{n-1}. \quad (5.52)$$

Таким образом, задачу Коши (5.48), (5.49) для дифференциального уравнения n -го порядка мы свели к задаче Коши для системы n уравнений первого порядка. К системе (5.51) можно применять рассмотренные в предыдущих подразделах численные методы.

5.4. Решение краевой задачи. Конечно-разностный метод

Рассмотрим построение разностной схемы на примере краевой задачи, которая весьма часто встречается в разнообразных приложениях:

$$y'' + p(x)y' + q(x)y = f(x), \quad x \in [0, 1] \quad (5.53)$$

$$y(0) = y_0, \quad y(1) = y_n. \quad (5.54)$$

Выражение (5.53) – неоднородное дифференциальное уравнение второго порядка, где $p(x)$, $q(x)$ и $f(x)$ – заданные функции, y_0 и y_n – заданные числа. Наряду с простыми условиями (5.54), позже мы рассмотрим более общий вид граничных условий. Существует преобразование, которое позволяет упростить уравнение (5.53) так, что в нем не будет члена с первой производной, однако мы рассмотрим более общий случай.

Разобьем отрезок $[0, 1]$ на n равных частей, введя сетку x_i ($i = 1, 2, \dots, n$). Вместо непрерывной функции $y(x)$, являющейся решением краевой задачи, введем сеточную функцию, определенную в узлах сетки: $y_i = y(x_i)$. Обозначим шаг сетки через h ($h = 1/n$). Возьмем произвольную точку сетки x_i , лежащую внутри интервала $[0, 1]$, и запишем в ней условия того, что в ней приближенно выполняется дифференциальное уравнение (5.53). Для этого рассмотрим две соседние с x_i точки x_{i-1} и x_{i+1} . Пусть $y_i = y(x_i)$, $y_{i-1} = y(x_{i-1})$, $y_{i+1} = y(x_{i+1})$. Запишем в точке x_i вместо y' и y'' их конечно-разностные формулы, использующие центральные разности (раздел 3):

$$y'_i = \frac{y_{i+1} - y_{i-1}}{2h} + O(h^2), \quad y''_i = \frac{y_{i+1} - 2y_i + y_{i-1}}{h^2} + O(h^2).$$

Тогда вместо уравнения (5.53) в точке x_i запишем его конечно-

разностную аппроксимацию

$$\frac{y_{i+1} - 2y_i + y_{i-1}}{h^2} + p_i \frac{y_{i+1} - y_{i-1}}{2h} + q_i y_i = f_i + O(h^2), \quad (5.55)$$

где $p_i = p(x_i)$, $q_i = q(x_i)$, $f_i = f(x_i)$. Так как $p(x)$, $q(x)$ и $f(x)$ – заданные функции, то p_i , q_i и f_i – постоянные величины для каждой конкретной точки x_i . Величина $h = \text{const}$ для каждого конкретного набора узловых точек. Если пренебречь малым членом $O(h^2)$, который определяет погрешность аппроксимаций, то (5.55) представляет собой линейное уравнение относительно трех неизвестных y_{i-1}, y_i, y_{i+1} . Уравнение вида (5.55) можно записать в каждом внутреннем узле сетки $x_1, x_2, \dots, x_{n-1}$. Количество таких уравнений равно количеству внутренних точек $n - 1$. Количество неизвестных в этих уравнениях $y_1, y_2, \dots, y_{n-1}$ также равно $n - 1$, так как в граничных точках x_0 и x_n ($x_0 = 0, x_n = 1$) значения функции $y(x)$ известны из граничных условий (5.54): $y_0 = y(x_0) = y(0) = y_0$, $y_n = y(x_n) = y(1) = y_n$. Если умножить обе части (5.55) на h^2 и собрать все члены при неизвестных, то систему линейных уравнений для всех внутренних точек можно записать в виде

$$\left(1 - \frac{p_i h}{2}\right) y_{i-1} + (q_i h^2 - 2) y_i + \left(1 + \frac{p_i h}{2}\right) y_{i+1} = f_i h^2 \quad (i = 1, 2, \dots, n-1).$$

Введя обозначения $a_i = 1 - \frac{p_i h}{2}$, $b_i = q_i h^2 - 2$, $c_i = 1 + \frac{p_i h}{2}$, $d_i = f_i h^2$,

эту систему можно записать в виде

$$a_i y_{i-1} + b_i y_i + c_i y_{i+1} = d_i \quad (i = 1, 2, \dots, n-1). \quad (5.56)$$

Это система разностных уравнений с переменными коэффициентами. Найдя все значения y_i из этой системы, мы получим приближенное решение нашей краевой задачи (5.53), (5.54). Систему (5.56) можно записать в компактном матричном виде

$$A \bar{y} = \bar{d}, \quad (5.57)$$

где матрица A , искомым вектор $\bar{y}$ и вектор правых частей $\bar{d}$

определены следующим образом:

$$A = \begin{bmatrix} b_1 & c_1 & 0 & 0 & & 0 & 0 & 0 \\ a_2 & b_2 & c_2 & 0 & \dots & 0 & 0 & 0 \\ 0 & a_3 & b_3 & c_3 & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & a_{n-2} & b_{n-2} & c_{n-1} \\ 0 & 0 & 0 & 0 & \dots & 0 & a_{n-1} & b_{n-1} \end{bmatrix} \quad \bar{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_{n-2} \\ y_{n-1} \end{bmatrix} \quad \bar{d} = \begin{bmatrix} d_1^* \\ d_2^* \\ d_3^* \\ \vdots \\ d_{n-2}^* \\ d_{n-1}^* \end{bmatrix} \quad (5.58)$$

Матрица A является трехдиагональной: во все уравнения, кроме первого и последнего, входят три искомых переменных y_{i-1} , y_i и y_{i+1} . Поясним, как вычисляются правые части d_1^* и d_{n-1}^* в первом и последнем уравнениях. При $i = 1$ из (5.56) следует, что

$$a_1 y_0 + b_1 y_1 + c_1 y_2 = d_1,$$

откуда, учитывая, что y_0 – число, задаваемое в граничном условии на левом конце интервала $[0, 1]$, получим

$$b_1 y_1 + c_1 y_2 = d_1 - a_1 y_0 = d_1^* \Rightarrow d_1^* = d_1 - a_1 y_0. \quad (5.59)$$

Аналогично в последнем уравнении (при $i = n - 1$)

$$a_{n-1} y_{n-2} + b_{n-1} y_{n-1} + c_{n-1} y_n = d_{n-1}$$

и, учитывая правое граничное условие, получим

$$a_{n-1} y_{n-2} + b_{n-1} y_{n-1} = d_{n-1} - c_{n-1} y_n \Rightarrow d_{n-1}^* = d_{n-1} - c_{n-1} y_n. \quad (5.60)$$

Рассмотрим конкретный пример записи разностной схемы для краевой задачи.

Пример 15. Найти численное решение дифференциального уравнения

$$y'' + xy' + y = 3x^2 + 2, \quad x \in [0, 1] \quad (5.61)$$

с краевыми условиями на отрезке $[0, 1]$

$$y(0) = 0, \quad y(1) = 1. \quad (5.62)$$

Решение. Решим задачу конечно-разностным методом. Разобьем отрезок $[0, 1]$ на 5 равных частей с шагом сетки, равным $0,2$. Не пользуясь общим видом записи системы уравнений (5.55) или (5.56), проведем для данного случая вывод этой системы уравнений. При $x = 0,2$ разностная аппроксимация уравнения имеет вид

$$\frac{y_2 - 2y_1 + y_0}{(0,2)^2} + 0,2 \frac{y_2 - y_0}{2 \cdot 0,2} + y_1 = 3 \cdot (0,2)^2 + 2.$$

Учитывая, что из левого граничного условия следует, что $y_0 = 0$, и проведя простые преобразования, получим

$$-49y_1 + 25,5y_2 = 2,12.$$

При $x = 0,4$ разностная аппроксимация дифференциального уравнения имеет вид

$$\frac{y_3 - 2y_2 + y_1}{(0,2)^2} + 0,4 \frac{y_3 - y_1}{2 \cdot 0,2} + y_2 = 3 \cdot (0,4)^2 + 2,$$

откуда после простых преобразований получим

$$24y_1 - 49y_2 + 26y_3 = 2,48.$$

При $x = 0,6$

$$\frac{y_4 - 2y_3 + y_2}{(0,2)^2} + 0,6 \frac{y_4 - y_2}{2 \cdot 0,2} + y_3 = 3 \cdot (0,6)^2 + 2,$$

откуда

$$23,5y_2 - 49y_3 + 26,5y_4 = 3,08.$$

При $x = 0,8$

$$\frac{y_5 - 2y_4 + y_3}{(0,2)^2} + 0,8 \frac{y_5 - y_3}{2 \cdot 0,2} + y_4 = 3 \cdot (0,8)^2 + 2.$$

Учитывая граничное условие на правом конце $y_5 = 1$, после преобразований получим

$$24y_3 - 49y_4 = -23,08.$$

Выпишем окончательную систему линейных уравнений для неизвестных величин y_1, y_2, y_3, y_4 :

$$\begin{cases} -49y_1 + 25,5y_2 & = 2,12 \\ 24y_1 - 49y_2 + 26y_3 & = 2,48 \\ 23,5y_2 - 49y_3 + 26,5y_4 & = 3,08 \\ 23y_3 - 49y_4 & = -23,08 \end{cases} \quad (5.63)$$

Используя для системы (5.63) метод Гаусса, получим решение, к которому для полноты добавлены граничные значения y_0 и y_5 :

$$y_0 = 0; y_1 = 0,04; y_2 = 0,16; y_3 = 0,36; y_4 = 0,64; y_5 = 1. \quad (5.63a)$$

Точное решение краевой задачи (5.61), (5.62) имеет вид $y = x^2$. В этом легко убедиться, подставив $y = x^2$ в дифференциальное уравнение (5.61) и учтя граничные условия (5.62). Интересно отметить, что разностные уравнения (5.63) также имеют решение $y = x^2$, в чем легко убедиться из вида решения (5.63a). В данном случае это произошло потому, что погрешности аппроксимации для первой производной с помощью центральной разности [формула (3.17)] и второй производной (формула (3.19)), которые использованы при записи разностной схемы для уравнения (5.61), равны нулю, так как для функции $y = x^2$ $y''(x) = 0$ и $y^{(4)}(x) = 0$, а эти производные входят в выражения для погрешностей в формулах (3.17) и (3.19).

Таким образом, конечно-разностный метод сводит решение краевой задачи для обыкновенного дифференциального уравнения к более простой задаче решения системы линейных алгебраических уравнений.

До сих пор мы рассматривали простой вид граничных условий (5.54), когда на границах интервала заданы значения решения $y(x)$. На практике часто встречаются граничные условия более общего вида:

$$\alpha_1 y'(0) + \beta_1 y(0) = \gamma_1; \quad (5.64)$$

$$\alpha_2 y'(1) + \beta_2 y(1) = \gamma_2. \quad (5.65)$$

В качестве частного случая этих общих условий могут быть заданы более простые граничные условия. Например, при $\alpha_1 = 0, \beta_1 \neq 0$ для

левой границы интервала получим уже рассмотренное условие

$$y(0) = \gamma_1 / \beta_1. \quad (5.66)$$

При $\alpha_1 = 1$, $\beta_1 = 0$ получим условие, в котором на границе задано значение производной от решения

$$y'(0) = \gamma_1. \quad (5.67)$$

Аналогичные условия могут быть заданы на правом конце. Рассмотрим, как записывать разностную схему в случае граничных условий общего вида (5.64) и (5.65). Система уравнений для внутренних точек не изменится и имеет вид (5.55). Граничные условия (5.64) и (5.65) также запишем в разностном виде, используя для $y'(0)$ и $y'(1)$ какие-либо конечно-разностные формулы. Например, используя односторонние разности (правую разность для точки $x = 0$ и левую для $x = 1$), заменим дифференциальные соотношения (5.64) и (5.65) разностными соотношениями:

$$\alpha_1 \frac{y_1 - y_0}{h} + \beta_1 y_0 = \gamma_1;$$

$$\alpha_2 \frac{y_n - y_{n-1}}{h} + \beta_2 y_n = \gamma_2,$$

или, приведя подобные члены и умножив обе части на h :

$$\alpha_1 y_1 + (\beta_1 h - \alpha_1) y_0 = \gamma_1 h; \quad (5.68)$$

$$(\alpha_2 + \beta_2 h) y_n - \alpha_2 y_{n-1} = \gamma_2 h. \quad (5.69)$$

Поскольку $\alpha_1, \beta_1, \gamma_1, \alpha_2, \beta_2, \gamma_2$ и h – заданные числа, формулы (5.68) и (5.69) являются линейными уравнениями относительно неизвестных y_0, y_1 и y_{n-1}, y_n . В отличие от предыдущего случая, теперь количество неизвестных, подлежащих определению, равно $n + 1$: $y_0, y_1, y_2, \dots, y_n$. Используя, как и раньше, во внутренних точках систему уравнений (5.55), которая состоит из $(n - 1)$ уравнения, добавим к ним уравнения (5.68) и (5.69) в граничных точках. В результате получим систему $(n + 1)$ линейных уравнений с таким же количеством неизвестных. Решив эту систему, мы найдем решение нашей краевой задачи (5.53), (5.68), (5.69).

Пример 16. Написать разностную схему для следующей краевой задачи

$$y'' - y = x, \quad y'(0) = 0, \quad y'(1) = 0, \quad x \in [0, 1], \quad (5.70)$$

взяв $n = 3$ ($h = 1/3$).

Решение. Возьмем сетку, состоящую из узловых точек $x_0 = 0, x_1 = 1/3, x_2 = 2/3, x_3 = 1$. Обозначим значения решений в этих точках y_0, y_1, y_2, y_3 . Запишем разностные уравнения во внутренних точках x_1 и x_2 :

$$\frac{y_2 - 2y_1 + y_0}{(1/3)^2} - y_1 = \frac{1}{3};$$

$$\frac{y_3 - 2y_2 + y_1}{(1/3)^2} - y_2 = \frac{2}{3}.$$

Для граничных точек используем односторонние разности:

$$\frac{y_1 - y_0}{(1/3)} = 0, \quad \frac{y_3 - y_2}{(1/3)} = 0.$$

После простых преобразований получим линейную систему относительно четырех неизвестных y_0, y_1, y_2, y_3 :

$$\begin{cases} y_0 - y_1 & & & = 0 \\ 9y_0 - 19y_1 + 9y_2 & & & = 1/3 \\ & 9y_1 - 19y_2 + 9y_3 & & = 2/3 \\ & & y_2 - y_3 & = 0 \end{cases} \quad (5.71)$$

В данном случае, выразив из первого уравнения $y_0 = y_1$, а из последнего $y_3 = y_2$ и подставив во второе и третье уравнения, мы получим систему двух уравнений для неизвестных y_1 и y_2 :

$$\begin{cases} 9y_1 - 19y_1 + 9y_2 & = 1/3 \\ & 9y_1 - 19y_2 + 9y_2 & = 2/3 \end{cases} \Rightarrow \begin{cases} -10y_1 + 9y_2 & = 1/3 \\ & 9y_1 - 10y_2 & = 2/3 \end{cases}$$

Система уравнений (5.71) является разностной схемой для краевой задачи (5.70).

Сделаем некоторые замечания относительно рассмотренных выше разностных схем для решения дифференциального уравнения (5.53).

Замечание 1. Граничные разностные условия (5.68) и (5.69) имеют первый порядок точности $O(h)$, так как для аппроксимации производных $y'(0)$ и $y'(1)$ были использованы односторонние разности [формулы (3.13) и (3.14)]. Используя формулы (3.27) и (3.29), можно записать аппроксимацию граничных условий со вторым порядком точности.

В случае граничных условий (5.64) и (5.65), используя (3.27) и (3.29), получим следующие разностные аппроксимации граничных условий:

$$\alpha_1 \left(\frac{-3y_0 + 4y_1 - y_2}{2h} \right) + \beta_1 y_0 = \gamma_1; \quad (5.72)$$

$$\alpha_2 \left(\frac{3y_n - 4y_{n-1} + y_{n-2}}{2h} \right) + \beta_2 y_n = \gamma_2, \quad (5.73)$$

которые будут иметь второй порядок точности.

Замечание 2. В подразд. 5.2 были указаны условия, при которых решение разностных уравнений сходится к решению дифференциального уравнения с граничными условиями. Для этого достаточно, чтобы разностные уравнения были устойчивы и аппроксимировали исходную дифференциальную задачу. Приведем без доказательства достаточные условия сходимости краевой разностной задачи.

$$\begin{cases} a_i y_{i-1} + b_i y_i + c_i y_{i+1} = d_i & (i = 1, 2, \dots, n-1) \\ y_0 = p, \quad y_n = q \end{cases} \quad (5.74)$$

где p и q – числа, определяющие граничные значения решения разностных уравнений.

Теорема. Если коэффициенты a_i, b_i, c_i разностной системы уравнений (5.74) ограничены в совокупности, т.е. если они удовлетворяют условиям

$$|a_i| < M, \quad |b_i| < M, \quad |c_i| < M, \quad (5.75)$$

где M – константа, не зависящая от i , и, кроме того, выполнены условия

$$|b_i| \geq |a_i| + |c_i| + \delta, \quad \delta > 0, \quad (5.76)$$

то решение разностной задачи $\{y_i\}$ будет единственным, и оно сходится к решению дифференциальной задачи, причем решение разностной задачи $\{y_i\}$ удовлетворяет оценке

$$|y_i| \leq \max \left\{ p, q, \frac{1}{\delta} \max_m |d_m| \right\}. \quad (5.77)$$

6. МЕТОДЫ РЕШЕНИЯ УРАВНЕНИЙ В ЧАСТНЫХ ПРОИЗВОДНЫХ

6.1. Основные понятия и постановка задачи

Большинство задач, описывающих сложные процессы, встречающиеся в природе, такие как движение жидкостей и газов, распространение тепла, механика твердых тел, процессы электродинамики и распространения электромагнитных волн и многие другие, описываются **уравнениями с частными производными**. Так называются уравнения, в которых искомая функция зависит от нескольких переменных, и уравнения содержат частные производные искомых функций. Как и в случае обыкновенных дифференциальных уравнений, уравнение с частными производными (или система таких уравнений) имеет бесчисленное множество решений. Для получения единственного решения, имеющего конкретный физический смысл, необходимо, помимо уравнения, задать некоторую дополнительную информацию. **Постановка задачи для уравнения в частных производных** включает вид уравнения, область решения задачи и дополнительные условия, обеспечивающие единственность решения задачи. Независимые переменные, от которых зависит искомая функция, могут быть пространственными координатами x, y, z , временем t и, быть может, некоторыми другими переменными. Если решение задачи не зависит от времени t , то задача называется *стационарной*, в противном случае – *нестационарной*. Если решение задачи ищется в ограниченной пространственной области, то обычно задаются условия на ее границе, называемые *граничными (краевыми) условиями*. Такие задачи называются **краевыми задачами** для уравнений в частных производных. Если одной из независимых переменных является время, то должны быть заданы начальные условия: значение искомого решения в начальный момент времени при $t = t_0$. Задача, которая состоит в

решении уравнения при заданных начальных условиях, называется **задачей Коши** для уравнения с частными производными, если задача решается в неограниченном пространстве и граничные условия не задаются. Нестационарные задачи, решаемые в ограниченной пространственной области, при формулировке которых ставятся и начальные, и граничные условия, называются смешанными.

Очень важным является вопрос о корректности постановки задачи. Задача называется **корректно поставленной**, если ее решение существует и единственно в некотором классе начальных и граничных условий и непрерывно зависит как от этих условий, так и от коэффициентов уравнения. Последнее означает, что для корректно поставленной задачи малому изменению коэффициентов уравнения и начальных и граничных условий соответствует малое изменение решения.

Свойства решений уравнений в частных производных и особенность постановки начальных и граничных условий зависят от типа уравнений. Для одних уравнений решения будут весьма гладкими, а для других существуют разрывные решения. Поэтому прежде всего разберем возможные типы уравнений в частных производных. Классификация этих уравнений является в общем случае достаточно сложной, и мы рассмотрим основные типы уравнений в случае двух независимых переменных. В этом случае уравнение можно записать в виде

$$a \frac{\partial^2 u}{\partial x^2} + 2b \frac{\partial^2 u}{\partial x \partial y} + c \frac{\partial^2 u}{\partial y^2} + d \frac{\partial u}{\partial x} + e \frac{\partial u}{\partial y} + fu = g. \quad (6.1)$$

Здесь x, y – независимые переменные, $u(x, y)$ – искомая функция, коэффициенты a, b, c, d, e, f и правая часть g в общем случае могут зависеть от переменных x, y и искомой функции $u(x, y)$. Если коэффициенты не зависят от x, y и решения $u(x, y)$, то уравнение (6.1) называют *уравнением с постоянными коэффициентами*. Если они зависят от x, y и не зависят от u , то уравнение называют *линейным*. Если все или некоторые из коэффициентов зависят не только от x, y , но и от функции $u(x, y)$ и некоторых ее производных, то уравнение называется *нелинейным*. В частном случае, когда коэффициенты зависят от $u(x, y)$, но не зависят от производных этой функции, уравнение называется *квазилинейным*.

В зависимости от соотношений между коэффициентами существуют различные виды уравнений. Напомним, что порядком уравнения называется порядок самой старшей производной, входящей в

уравнение. Перечислим некоторые виды уравнений, являющихся частным случаем общего уравнения (6.1).

При $a = b = c = f = 0$, $d \neq 0$, $e \neq 0$ получается уравнение первого порядка, часто называемое уравнением переноса

$$\frac{\partial u}{\partial x} + S \frac{\partial u}{\partial y} = q. \quad (6.2)$$

В качестве одной из независимых переменных часто используется время. Тогда уравнение (6.2) (или система подобных уравнений) называется *эволюционным* и описывает распространение во времени и в пространстве какой-либо субстанции: массы, количества движения, теплоты и т.п.

Если хотя бы один из коэффициентов a , b , c , отличен от нуля, то уравнение (6.1) будет уравнением второго порядка. В зависимости от знака дискриминанта $D = b^2 - 4ac$ оно может принадлежать к одному из трех типов: *гиперболическому* ($D > 0$), *параболическому* ($D = 0$) или *эллиптическому* ($D < 0$). На практике встречаются более сложные задачи, когда в разных областях решения величина D может иметь разные знаки. Такие задачи называют *смешанными*. Примерами рассмотренных типов задач являются:

– гиперболическое волновое уравнение

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2}, \quad (u = u(x, t), \quad a > 0); \quad (6.3)$$

– параболическое уравнение нестационарной теплопроводности

$$\frac{\partial u}{\partial t} = a \frac{\partial^2 u}{\partial x^2}, \quad (u = u(x, t), \quad a > 0); \quad (6.4)$$

– эллиптическое уравнение Пуассона

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = f(x, y), \quad (u = u(x, y)). \quad (6.5)$$

При $f(x, y) = 0$ уравнение (6.5) называется уравнением Лапласа

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0. \quad (6.6)$$

Уравнения (6.3) – (6.6) часто называют уравнениями математической физики, так как они описывают очень многие физические процессы. В реальных процессах, которые моделируются такими уравнениями, приходится использовать больше независимых переменных, и уравнения имеют более сложный вид, однако принципиальные основные особенности этих типов уравнений сохраняются. Уравнения, решения которых зависят от одной пространственной независимой переменной (например, x), называют одномерными, от двух переменных – двумерными, от трех переменных – трехмерными. Уравнения (6.3) и (6.4) будут одномерными нестационарными, а уравнения (6.5) и (6.6) – двумерными стационарными. Обобщением уравнения (6.4) на трехмерный случай является трехмерное нестационарное уравнение теплопроводности

$$\frac{\partial u}{\partial t} = a \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} \right) \quad (u = u(x, y, z, t), \quad a > 0). \quad (6.7)$$

Для некоторых ограниченных классов уравнений математической физики можно получить аналитические решения. Это, как правило, уравнения с постоянными коэффициентами и с простой формой геометрической области определения независимых переменных, в которой решается задача (круг, квадрат, прямоугольник и т.п.). Однако часто даже в этом случае решение имеет вид сходящихся рядов довольно сложной структуры, а коэффициенты, от которых зависят эти ряды, определяются путем решения одного или нескольких нелинейных алгебраических или трансцендентных уравнений. Типичным примером таких методов является метод разделения переменных Фурье для задач (6.3) – (6.6). Для получения конечных численных результатов с помощью таких аналитических методов приходится использовать численные методы для суммирования рядов и для решения задач на собственные значения, причем для вычислений, как правило, используются компьютеры. Однако получение аналитических решений имеет большое значение, и не только для того, чтобы использовать их затем для решения различных практических задач. Их исследование позволяет развивать общую теорию и исследовать общие принципиальные вопросы, которые затем используются при численном моделировании более сложных задач, не имеющих аналитических решений. Кроме того, аналитические решения широко используются для проверки с их помощью различных численных методов решения сложных задач и для отладки различных создаваемых комплексов программ для ЭВМ. К счастью, такие *модельные задачи*, для которых существует аналитическое решение, имеются не только для линейных уравнений математической физики,

но и для некоторых весьма сложных нелинейных задач. Например, для таких, как система нелинейных уравнений, описывающих движение жидкости или газа.

Существует большое количество разнообразных численных методов решения уравнений с частными производными. Наиболее часто используемые методы: вариационные, прямые, конечно-разностные (метод сеток).

Главная идея вариационных методов – переход от задачи для уравнения с частными производными к задаче минимизации специальным образом построенного функционала. Напомним, что функционалом называется такой оператор (отображение), который каждой функции ставит в соответствие некоторое число. Простым примером функционала является определенный интеграл

$$S = \int_a^b f(x)dx,$$

который каждой допустимой функции (подынтегральной функции) ставит в соответствие число – величину определенного интеграла.

Прямые методы основаны на представлении решения задачи в виде суммы линейной комбинации некоторой базисной системы функций. Коэффициенты этой линейной комбинации определяются из условия наилучшей в некотором смысле аппроксимации решения краевой задачи или задачи Коши для решаемого уравнения с частными производными.

Главная идея метода сеток, как и в обыкновенных дифференциальных уравнениях, – переход от непрерывной функции, являющейся решением уравнения с соответствующими дополнительными условиями, к дискретной сеточной функции, определенной в узлах сетки и приближенно аппроксимирующей решение задачи.

Во всех трех типах методов линейные задачи, в конечном счете, сводятся к решению систем линейных уравнений высокого порядка с матрицами специального вида, имеющими много нулевых элементов. Однако способ получения линейных систем у разных методов является различным. Для нелинейных уравнений в частных производных задача сводится к решению системы нелинейных алгебраических уравнений, которые решаются какими-либо итерационными методами. Учитывая, что вариационные и прямые методы требуют хорошего знания специальных разделов математики, мы в данном пособии их рассматривать не будем. Хорошее изложение этих методов приведено, например, в работах [3, 5]. В этом разделе мы рассмотрим метод сеток, который, пожалуй, является наиболее универсальным методом, применимым и для линейных, и для нелинейных уравнений с частными

производными с разнообразными граничными условиями, которые также могут быть как линейными, так и нелинейными.

6.2. Описание метода сеток. Аппроксимация, устойчивость, сходимость

Метод сеток – это способ решения уравнений в частных производных. Он основан на:

- рассмотрении дискретной области независимых переменных (сетки), в которой решается задача;
- выражении искомой функции, ее производных, а также начальных и граничных условий через значения этой функции в узлах сетки.

В результате таких преобразований получается система алгебраических уравнений, называемая разностной схемой.

При использовании метода сеток мы от непрерывной функции переходим к сеточной, определенной в узлах сетки. Поэтому прежде всего заменим область независимых переменных, в которой ищется решение задачи, сеточной областью. Пусть, например, область решения задачи $G(x,y)$ представляет собой прямоугольник $G = \{a \leq x \leq b, c \leq y \leq d\}$ (рис 6.1). Разделим отрезок $[a, b]$ на n равных частей, введя узловые точки $x_0 = a, x_1, x_2, \dots, x_n = b$, расположенные на одинаковом

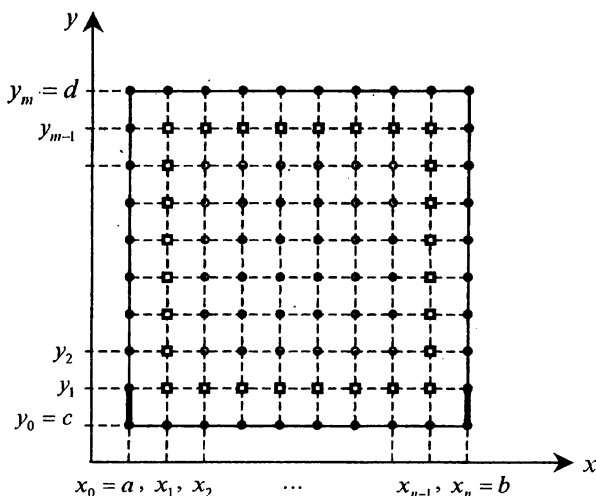


Рис 6.1. Сеточная область решения задачи

(• – граничные узлы, ◻ – приграничные узлы, ○ – внутренние узлы)

расстоянии Δx друг от друга. Общая формула для произвольного узла по оси x $x_i = a + i\Delta x$. Разделив отрезок $[c, d]$ на m равных частей, введем узловые точки $y_0 = c, y_1, y_2, \dots, y_m = d$, общая формула для которых будет $y_j = c + j\Delta y$, где Δy – шаг сетки по оси y . Шаги сетки Δx и Δy связаны с количеством интервалов n и m очевидными соотношениями: $\Delta x = (b - a)/n$, $\Delta y = (d - c)/m$. Через точки x_i и y_j проводятся прямые, параллельные осям x и y , которые образуют сетку, узлами которой будут точки (x_i, y_j) . Положение каждого узла можно определить целочисленными координатами (i, j) . Узлы сетки, лежащие на границе Γ области G , называют граничными узлами. На рис. 6.1 граничные узлы изображены черными кружками. Остальные узлы сетки являются внутренними. Внутренние узлы, которые имеют рядом хотя бы один граничный узел, назовем приграничными узлами. Для некоторых задач алгоритм расчета в приграничных точках отличается от алгоритма расчета во внутренних точках, не являющихся приграничными.

Для простоты мы рассматриваем прямоугольные сетки, считая, что область задана в декартовой системе координат. В реальных практических задачах иногда приходится использовать и другие системы координат: полярную, сферическую или цилиндрическую. В этом случае и уравнение, и дополнительные условия, и область решения задачи записываются в наиболее удобной для решения задачи системе координат. Иногда в расчетной сеточной области используют неравномерную сетку. Это необходимо в том случае, когда имеются области, в которых существуют большие градиенты искомого решения. В этом случае для сохранения точности расчета необходимо сгущение сетки в области больших градиентов.

При построении разностной схемы частные производные в уравнениях заменяются конечно-разностными соотношениями. В разделе 3 были получены различные формулы, аппроксимирующие производные для функции одной переменной. Приведем аналогичные формулы для функции многих переменных. Пусть для простоты задана функция двух переменных $u(x, y)$ и введена сетка, изображенная на рис. 6.1.

Напомним, что частной производной $\frac{\partial u(x, y)}{\partial x}$ называется производная

по x при фиксированном значении y , а производной $\frac{\partial u(x, y)}{\partial y}$ называется производная по переменной y при фиксированном значении x .

Например, производная $\frac{\partial u}{\partial x}$ в данной точке (x_0, y_0) определяется как

$$\left(\frac{\partial u}{\partial x} \right)_{x_0, y_0} = \lim_{\Delta x \rightarrow 0} \frac{u(x_0 + \Delta x, y_0) - u(x_0, y_0)}{\Delta x}.$$

Обозначим значение функции в узловой точке через $u_{i,j} = u(x_i, y_j)$. Тогда, используя формулы (3.13), (3.14) и (3.17), напомним конечно-разностные формулы для первой производной $\frac{\partial u}{\partial x}$:

$$\left(\frac{\partial u}{\partial x} \right)_{i,j} = \frac{u_{i+1,j} - u_{i,j}}{\Delta x} + O(\Delta x); \quad (6.8)$$

$$\left(\frac{\partial u}{\partial x} \right)_{i,j} = \frac{u_{i,j} - u_{i-1,j}}{\Delta x} + O(\Delta x); \quad (6.9)$$

$$\left(\frac{\partial u}{\partial x} \right)_{i,j} = \frac{u_{i+1,j} - u_{i-1,j}}{2\Delta x} + O(\Delta x^2). \quad (6.10)$$

Обозначение $\left(\frac{\partial u}{\partial x} \right)_{i,j}$ означает производную $\frac{\partial u(x_i, y_j)}{\partial x}$ в точке (x_i, y_j) . Формула (6.8) использует правую разность, (6.9) – левую, а (6.10) – центральную разность; первые две имеют первый порядок точности, а последняя – второй порядок точности. Поясним, например, формулу (6.8). Изменение первого индекса i означает, что берется производная по x , второй индекс j остается неизменным, так как при вычислении частной производной по x переменная y остается неизменной. Для второй производной, используя формулу (3.19), получим

$$\left(\frac{\partial^2 u}{\partial x^2} \right)_{i,j} = \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{(\Delta x)^2} + O(\Delta x^2). \quad (6.11)$$

Аналогичные формулы можно записать для частных производных по y , причем теперь второй индекс будет переменным, а первый –

ПОСТОЯННЫМ:

$$\left(\frac{\partial u}{\partial y}\right)_{i,j} = \frac{u_{i,j+1} - u_{i,j}}{\Delta y} + O(\Delta y); \quad \left(\frac{\partial u}{\partial y}\right)_{i,j} = \frac{u_{i,j} - u_{i,j-1}}{\Delta y} + O(\Delta y); \quad (6.12)$$

$$\left(\frac{\partial u}{\partial y}\right)_{i,j} = \frac{u_{i,j+1} - u_{i,j-1}}{2\Delta y} + O(\Delta y^2); \quad (6.13)$$

$$\left(\frac{\partial^2 u}{\partial y^2}\right)_{i,j} = \frac{u_{i,j+1} - 2u_{i,j} + u_{i,j-1}}{(\Delta y)^2} + O(\Delta y^2). \quad (6.14)$$

Используя приведенные формулы, можно написать конечно-разностную аппроксимацию уравнения в частных производных. Для этого каждую входящую в уравнение частную производную надо заменить подходящим конечно-разностным выражением. В следующих подразделах мы рассмотрим запись разностных уравнений Пуассона и теплопроводности и способы их решения. А сейчас разберем основные понятия теории разностных схем: аппроксимацию дифференциального выражения разностным, устойчивость разностной схемы и вопрос о сходимости решения разностного уравнения к решению дифференциального уравнения при уменьшении шага сетки. Эти вопросы уже обсуждались в разделе 5 применительно к обыкновенным дифференциальным уравнениям. Поэтому здесь будут кратко изложены основные понятия и результаты, необходимые для дальнейшего изложения.

Пусть $U(x, y)$ – решение какого-либо уравнения в частных производных с соответствующими начальными и граничными условиями. Запишем всю задачу в операторном виде

$$LU(x, t) = F(x, t), \quad (x, t) \in G. \quad (6.15)$$

Операторное уравнение включает в себя как исходное уравнение, так и дополнительные (начальные и граничные) условия. Область G является расчетной областью решения задачи вместе с ее границей. Заменим область G дискретной сеточной областью g_h , где h – параметр сетки. Для простоты предполагается, что шаг сетки по времени τ зависит от шага сетки h по оси x : $\tau = rh$, где $r = \text{const}$. Тогда разностную задачу можно также записать в операторном виде

$$L_h u_h = f_h, \quad (x, t) \in g_h, \quad (6.16)$$

где u_h – сеточная функция, являющаяся решением разностного уравнения. Рассмотрим погрешность решения в какой-либо точке сетки

$$\delta u_h = u_h - U.$$

Здесь U – точное решение дифференциальной задачи, а u_h – точное решение разностной задачи, δu_h – величина погрешности. Отсюда можно выразить

$$u_h = U + \delta u_h. \quad (6.17)$$

Подставив это выражение в (6.16), получим:

$$\begin{aligned} L_h U_h + L_h(\delta u_h) &= f_h; \\ L_h(\delta u_h) &= f_h - L_h U_h = R_h. \end{aligned} \quad (6.18)$$

Величина R_h называется *погрешностью аппроксимации разностной схемы* в данной точке. Введем характерную интегральную величину невязки, положив ее равной максимальному значению невязки

$$R = \max_{(x,t) \in g_h} |R_h|, \quad (6.19)$$

R – максимальная погрешность в сеточной области. Говорят, что *аппроксимация имеет порядок k , если $R = O(h^k)$* . Ранее мы считали, что шаги сетки по x и t , равные h и τ , связаны между собой соотношением $\tau = rh$. Если они независимы и имеет место соотношение $R = O(h^k + \tau^l)$, то говорят, что *аппроксимация имеет порядок k по h и порядок l по τ* .

Определение 1. Разностная схема (6.16) аппроксимирует исходную дифференциальную задачу (6.15), если при измельчении сетки по всем переменным погрешность аппроксимации стремится к нулю:

$$\lim R = 0 \quad (h \rightarrow 0, \tau \rightarrow 0). \quad (6.20)$$

Рассмотрим погрешность решения в произвольной точке $\delta u_h = u_h - U$ и максимальное по модулю значение этой погрешности на сетке

$$\delta u = \max |\delta u_h|, \quad (x, t) \in g_h. \quad (6.21)$$

Определение 2. Разностная схема (6.16) называется сходящейся, если при сгущении узлов сетки значение погрешности стремится к нулю

$$\lim \delta u = 0 \quad (h \rightarrow 0, \tau \rightarrow 0). \quad (6.22)$$

Определение 3. Разностная схема (6.16) называется устойчивой, если малому изменению входных данных (правой части и коэффициентов системы разностных уравнений) соответствует малое изменение решения.

Следует подчеркнуть, что устойчивость разностной схемы является ее внутренним свойством, которое никак не связано с исходной дифференциальной задачей (6.15). Устойчивость схемы характеризует ее чувствительность к различным погрешностям. Напротив, аппроксимация и сходимость существенно учитывают свойства исходной дифференциальной задачи.

Неустойчивость разностной схемы может быть условной и безусловной. Разностная схема называется *безусловно устойчивой*, если она устойчива при любых соотношениях между шагами сетки по независимым переменным (например, между h и τ). Разностная схема называется *условно устойчивой*, если она устойчива только при определенном соотношении между шагами сетки h и τ .

Фундаментальное значение в теории разностных схем имеет следующее утверждение.

Теорема. Если решение исходной дифференциальной задачи (6.15) существует, а разностная схема устойчива и аппроксимирует задачу (6.15) на данном решении, то решение разностного уравнения сходится к решению дифференциального уравнения.

Другими словами, из аппроксимации и устойчивости разностной схемы следует ее сходимость. В практическом плане это означает, что нам достаточно проверить, что разностная схема аппроксимирует исходную дифференциальную задачу и является устойчивой. Сходимость отсюда следует автоматически. Непосредственная проверка сходимости является достаточно трудной задачей.

Ниже все введенные понятия мы рассмотрим на конкретных примерах разностных схем для решения некоторых уравнений в частных производных.

6.3. Решение параболических задач

Рассмотрим решение этих задач на примере одномерного нестационарного уравнения теплопроводности

$$\frac{\partial U}{\partial t} = a \frac{\partial^2 U}{\partial x^2}, \quad U = U(x, t), \quad t > 0, \quad x \in [0, 1] \quad (6.23)$$

для смешанной краевой задачи. Начальные условия возьмем в виде

$$U(x,0) = \varphi(x) \text{ при } t = 0. \quad (6.24)$$

Граничные условия запишем в виде:

$$\begin{aligned} \text{при } x = 0 \quad U(0,t) &= U_0, \\ \text{при } x = 1 \quad U(1,t) &= U_1. \end{aligned} \quad (6.25)$$

Функция $U(x,t)$ зависит от двух переменных: x – пространственная координата, t – время. Если U является температурой, то задача (6.23) – (6.25) описывает распространение температуры на отрезке $0 \leq x \leq 1$ в различные моменты времени при постоянных заданных значениях температур на концах отрезка [граничные условия (6.25)] и заданном исходном распределении температур $\varphi(x)$ в начальный момент времени. Область решения задачи приведена на рис. 6.2. Эта область является прямоугольником $G = \{ 0 \leq x \leq 1, 0 \leq t \leq t_k \}$, где t_k – конечный момент времени, в котором надо получить решение задачи.

Введем равномерную прямоугольную сетку с помощью координатных линий $x_i = ih$ ($i = 0, 1, 2, \dots, n$) и $t_j = j\tau$ ($j = 0, 1, 2, \dots, k$), где h и

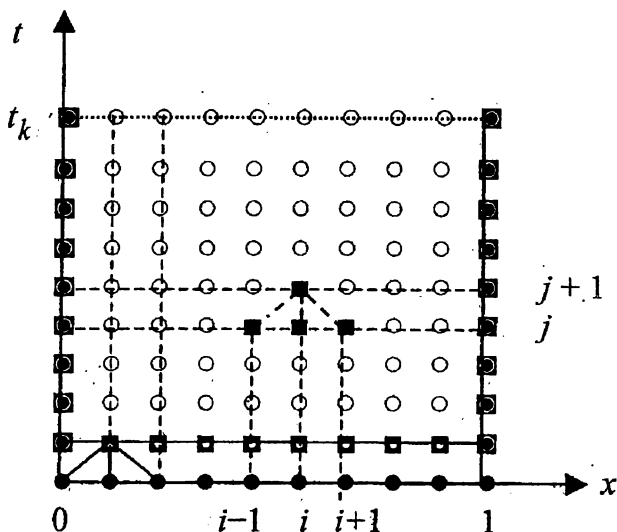


Рис.6.2

● – начальные условия, ◻ – граничные условия

τ – шаги сетки по x и t . Введем сеточную функцию $u(x_i, t_j) = u_{i,j}$, которую надо определить во всех внутренних узлах области. Получим систему разностных уравнений для уравнения (6.23). Для производных $\partial U / \partial t$ и $\partial^2 U / \partial x^2$ используем следующие конечно-разностные выражения [см. формулы (6.8) и (6.11)]:

$$\left(\frac{\partial U}{\partial t} \right)_{i,j} = \frac{u_{i,j+1} - u_{i,j}}{\tau} + O(\tau); \quad (6.26)$$

$$\left(\frac{\partial^2 U}{\partial x^2} \right)_{i,j} = \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h^2} + O(h^2). \quad (6.27)$$

Рассмотрим произвольную внутреннюю точку (i, j) сеточной области и запишем условие того, что в этой точке приближенно выполняется уравнение теплопроводности (6.23). Для этого производные, входящие в (6.23), заменим на их разностные выражения (6.26) и (6.27)

$$\frac{u_{i,j+1} - u_{i,j}}{\tau} + O(\tau) = a \left(\frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h^2} \right) + O(h^2). \quad (6.28)$$

Отбросив малые члены $O(\tau) + O(h^2)$, умножив обе части уравнения на τ , перенеся вправо член $u_{i,j}$ и приведя подобные члены, получим следующее разностное уравнение

$$u_{i,j+1} = u_{i+1,j} + \left(1 - \frac{2a\tau}{h^2} \right) u_{i,j} + u_{i-1,j}. \quad (6.29)$$

Схема (6.29) называется *явной разностной схемой*. В этой схеме используются значения сеточной функции на двух временных слоях j и $j + 1$, причем на слое $(j + 1)$ есть только одна сеточная функция $u_{i,j+1}$, а на слое j – три значения: $u_{i+1,j}$, $u_{i,j}$ и $u_{i-1,j}$. В разностных схемах часто используется понятие *шаблона разностной схемы*, на котором схематически показано, какие сеточные функции входят в разностную схему в произвольной точке (i, j) . Шаблон для явной схемы (6.29) приведен на рис. 6.3. Схема (6.29) называется явной, так как с ее помощью можно найти значение функции $u_{i,j+1}$ в следующий момент

времени $j + 1$, если известны значения функций в трех соседних узлах $u_{i+1,j}$, $u_{i,j}$ и $u_{i-1,j}$ в момент времени j .

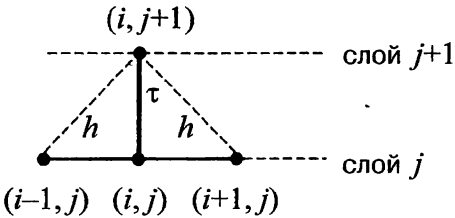


Рис 6.3

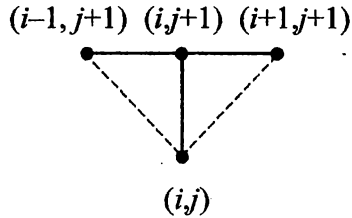


Рис. 6.4

Покажем, что пользуясь разностной схемой, мы можем найти решение задачи во всей сеточной области G (рис. 6.2). Начальные условия (6.24) для сеточной функции располагаются на слое $j = 0$ (точки этого слоя отмечены на рис. 6.2 черными кружками): $u_{i,0} = \varphi(x_i)$. Там же отмечены точки сетки, где заданы граничные условия: $u_{0,j} = U_0$, $u_{n,j} = U_1$. Используя начальные условия $u_{0,0}, u_{1,0}, u_{2,0}$ в первых трех точках $(0, 0)$, $(1, 0)$, $(2, 0)$ нулевого временного слоя, мы можем по явной схеме найти решение $u_{1,1}$. Сдвинувшись вправо на один шаг, таким же образом найдем $u_{2,1}$, затем $u_{3,1}$ и так до тех пор, пока не определим значения решения во всем временном слое $j = 1$. После этого, принимая за исходный временной слой $j = 1$, последовательно по явной схеме определим решение на слое $j = 2$. Таким образом, переходя от одного временного слоя j к следующему слою $j + 1$, мы последовательно определим решение во всей сеточной области G .

Как отмечалось в предыдущем подразделе, для практического использования явной схемы надо показать, что она аппроксимирует уравнение (6.23) и является устойчивой.

Аппроксимацию явной схемы в произвольной точке (i, j) запишем как

$$R_{i,j} = \frac{u_{i,j+1} - u_{i,j}}{\tau} - a \left(\frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h^2} \right).$$

Подставив вместо разностных соотношений их выражение через производные точного решения (6.26), (6.27), получим

$$R_{i,j} = \left[\left(\frac{\partial U}{\partial t} \right)_{i,j} - a \left(\frac{\partial^2 U}{\partial x^2} \right)_{i,j} \right] + O(\tau) + O(h^2).$$

Если учесть, что $U(x_i, t_j)$ является точным решением уравнения теплопроводности (6.23), то выражение в квадратных скобках равно нулю, откуда окончательно получим

$$R = \max_{i,j} |R_{i,j}| = O(\tau) + O(h^2). \quad (6.30)$$

Таким образом, явная разностная схема аппроксимирует уравнение (6.23) со вторым порядком точности по h и с первым порядком по временному шагу τ . Начальные и граничные условия аппроксимируются на границах точно, поскольку здесь значения сеточной функции равны соответствующим значениям точных граничных условий.

Исследование устойчивости схемы проводится более сложно. Поэтому мы его проводить не будем, а приведем лишь окончательный результат. *Явная схема является условно устойчивой.* Она устойчива лишь тогда, когда выбранные шаги по пространству и времени (h и τ) удовлетворяют неравенству

$$\frac{a\tau}{h^2} \leq \frac{1}{2}. \quad (6.31)$$

В противном случае явная схема неустойчива и ее для расчетов использовать нельзя. Из аппроксимации схемы и ее устойчивости следует ее сходимость со скоростью $O(\tau + h^2)$. На практике величину h выбирают так, чтобы она обеспечивала необходимую точность результатов, а τ выбирается как из условий необходимой точности, так и из условия устойчивости (6.31) $\tau \leq h^2/2a$. При больших значениях физического параметра a ограничение на τ является очень обременительным и приходится вести расчеты с очень мелкими шагами по времени, что требует много машинного времени. Были разработаны другие разностные схемы, которые свободны от этого недостатка. Рассмотрим одну из них – неявную разностную схему.

При аппроксимации члена $\partial^2 U / \partial x^2$ в уравнении (6.23) мы использовали аппроксимацию (6.27), которая использует значение

второй производной на исходном временном слое j . Но при переходе от слоя j к слою $j + 1$ величина $\partial^2 U / \partial x^2$ изменяется и по x , и по t , и вовсе не обязательно брать ее значение на слое j . В неявной схеме она берется на слое $j + 1$, решение на котором пока неизвестно и подлежит определению. Запишем неявную схему в виде

$$\frac{u_{i,j+1} - u_{i,j}}{\tau} = a \left(\frac{u_{i+1,j+1} - 2u_{i,j+1} + u_{i-1,j+1}}{h^2} \right). \quad (6.32)$$

Умножив обе части уравнения на τ , обозначив $\frac{a\tau}{h^2} = \lambda$ и перенеся все значения сеточной функции, подлежащие определению на слое $(j + 1)$, в левую сторону, получим

$$\lambda u_{i-1,j+1} - (1 + 2\lambda) u_{i,j+1} + \lambda u_{i+1,j+1} = u_{i,j}. \quad (6.33)$$

Шаблон для неявной схемы приведен на рис.6.4. Неявная схема (6.33) не позволяет по простой формуле найти значения сеточной функции в следующий момент времени $(j + 1)$ по значениям ее в предыдущий момент времени j , как это было в явной схеме. Если (6.33) записать в каждой точке (i, j) , то для определения решения мы получим систему линейных уравнений (6.33), для которой $i = 1, 2, 3, \dots, n - 1$. Количество неизвестных в этой системе и количество уравнений в ней равно количеству внутренних точек на слое $(j + 1)$ и, следовательно, составляет $n - 1$. Эта система уравнений имеет трехдиагональную матрицу и, следовательно, может быть легко решена методом Гаусса. Для таких матриц существует простой и эффективный способ реализации метода Гаусса, называемый методом прогонки. Метод прогонки очень широко используется в вычислительной математике во всех случаях, когда требуется решать систему линейных уравнений с трехдиагональной матрицей. Описание этого метода приведено, например, в работах [5, 11]. Поскольку по неявной схеме на каждом шаге по времени приходится решать систему линейных уравнений, она требует для получения решения больше операций, чем явная схема. Однако это неудобство компенсируется тем, что неявная схема является абсолютно устойчивой и, следовательно, не накладывает никаких дополнительных ограничений на шаги h и τ , кроме тех, которые обусловлены необходимой точностью расчета. Отсюда следует, что при больших значениях параметра a , входящего в уравнение (6.23), неявная

схема предпочтительней явной схемы. Как отмечалось выше, в явной схеме в этом случае приходится брать очень маленькие шаги по времени τ . И хотя неявная схема при расчете одного временного слоя требует примерно в четыре раза больше арифметических операций, чем явная, в ней приходится считать гораздо меньше временных шагов, что и определяет большую эффективность неявной схемы.

В заключение отметим, что для уравнения теплопроводности, помимо простых граничных условий (6.25), часто встречаются и более сложные граничные условия, когда на границе задана линейная комбинация решения и его производной. В этом случае необходимо аппроксимировать не только исходное дифференциальное уравнение, но и граничные условия. При этом в неявной схеме на шаге $(j+1)$ количество уравнений и искомых переменных будет на два больше.

6.4. Решение эллиптических задач

Эллиптические задачи обычно описывают стационарные режимы различных процессов. Разберем методы решения этих задач на примере уравнения Пуассона

$$\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} = f(x, y). \quad (6.34)$$

Здесь $U(x, y)$ – искомая функция, а $f(x, y)$ – заданная функция. Для эллиптических задач, помимо уравнения, необходимо задать область решения задачи и граничные условия на границе этой области. Для простоты в качестве области решения задачи возьмем квадрат единичной длины $G = \{ 0 \leq x \leq 1, 0 \leq y \leq 1 \}$ (рис. 6.5). Границей области Γ являются стороны этого квадрата.

Граничные условия в общем случае могут иметь один из следующих видов:

$$U(x, y)|_{\Gamma} = \varphi(x, y); \quad (6.35)$$

$$\left. \frac{\partial U}{\partial n} \right|_{\Gamma} = \varphi_1(x, y); \quad (6.36)$$

$$\left(\alpha \frac{\partial U}{\partial n} + \beta U \right) \Big|_{\Gamma} = \varphi_2(x, y). \quad (6.37)$$

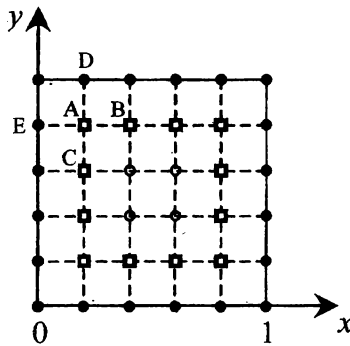


Рис. 6.5

Граничное условие (6.35) означает задание значений искомой функции на границах области, (6.36) – задание значения ее нормальной производной на границе, а (6.37) – их линейной комбинации. Условие (6.37) является наиболее общим. В частном случае из него получается условие (6.35) (при $\alpha = 0$, $\beta = 1$) и условие (6.36) (при $\alpha = 1$, $\beta = 0$). Условие (6.35) называют условием Дирихле, а соответствующую краевую задачу (6.34), (6.35) – задачей Дирихле. Условие (6.36) и соответствующую задачу называют задачей Неймана, а задачу с условием (6.37) – смешанной задачей.

Рассмотрим задачу Дирихле (6.34), (6.35) для квадрата, приведенного на рис. 6.5. Точки, где заданы граничные условия, отмечены черными кружками. В этом случае условие (6.35) можно записать в виде

$$U(x, y)|_G = \varphi(x, y) = \begin{cases} \varphi_1(y) & x=0 \\ \varphi_2(y) & x=1 \\ \varphi_3(x) & y=0 \\ \varphi_4(x) & y=1 \end{cases} \quad (6.38)$$

Введем в области G прямоугольную сетку $x_i = ih$ ($i = 0, 1, 2, \dots, n$), $y_j = jh$ ($j = 0, 1, 2, \dots, n$), узлы которой располагаются на пересечении прямых $x = x_i, y = y_j$. Точное значение решения задачи $U(x_i, y_j)$ заменим сеточной функцией $u_{i,j} = u(x_i, y_j)$. Запишем условие того, что в произвольной точке (x_i, y_j) приближенно выполняется исходное уравнение (6.34). Для этого заменим вторые производные $\partial^2 U / \partial x^2$ и $\partial^2 U / \partial y^2$ их конечно-разностными соотношениями (6.11) и (6.14)

$$\frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h^2} + \frac{u_{i,j+1} - 2u_{i,j} + u_{i,j-1}}{h^2} = f(x_i, y_j). \quad (6.39)$$

Формулы (6.11) и (6.14) показывают, что погрешность аппроксимации в этом случае имеет по h второй порядок $O(h^2)$. Умножив обе части уравнения на h^2 и обозначив $d_{i,j} = h^2 f(x_i, y_j)$, получим, приведя подобные члены

$$u_{i+1,j} + u_{i-1,j} + u_{i,j+1} + u_{i,j-1} - 4u_{i,j} = d_{i,j} \quad (6.40)$$

$$(i, j = 1, 2, \dots, n-1).$$

Уравнение (6.40) можно записать в каждой внутренней точке. Всего таких точек $(n-1)(n-1)$. Следовательно, (6.40) – система линейных уравнений порядка $(n-1)(n-1)$. Количество неизвестных также равно $(n-1)(n-1)$, так как неизвестными являются значения сеточной функции $u_{i,j}$ во внутренних точках, ибо на границе $u_{i,j}$ заданы граничными условиями (6.38), которые записываются точно и для разностных уравнений:

$$u_{0,j} = \varphi_1(y_j), \quad u_{n,j} = \varphi_2(y_j), \quad u_{i,0} = \varphi_3(x_i), \quad u_{i,n} = \varphi_4(x_i) \quad (6.41)$$

$$(i, j = 0, 1, 2, \dots, n).$$

Можно показать, что решение (6.40), (6.41) существует. Эта разностная схема аппроксимирует исходную дифференциальную задачу (6.34), (6.35) в общем случае с погрешностью $O(\Delta x^2) + O(\Delta y^2)$, где Δx и Δy – шаги сетки соответственно по x и по y , и является устойчивой.

Матрица линейной системы (6.40) является пятидиагональной: у нее не равны нулю элементы, стоящие на главной диагонали, на двух соседних диагоналях и еще на двух диагоналях, отстоящих от главной на $(n-1)$ (рис.6.6).

В каждое уравнение входят только пять ненулевых членов: на главной диагонали элемент матрицы равен (-4) и он соответствует неизвестной $u_{i,j}$, остальные элементы равны 1, они соответствуют неизвестным $u_{i-1,j}$ и $u_{i+1,j}$, расположенным рядом с $u_{i,j}$, и неизвестным $u_{i,j+1}$ и $u_{i,j-1}$.

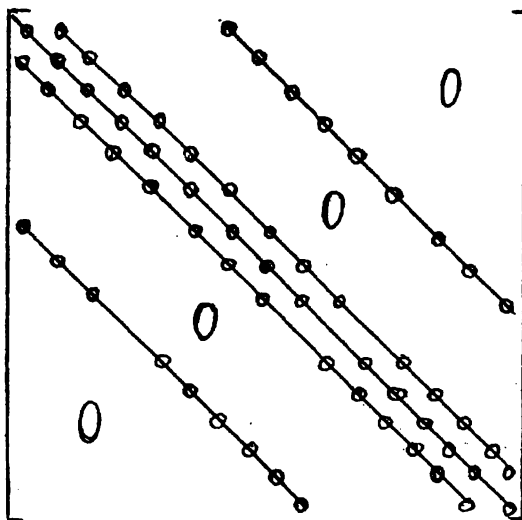


Рис. 6.6

Существует много различных способов решения системы линейных уравнений (6.40). Можно использовать как специальные прямые методы, так и итерационные. Большой набор таких методов для решения разнообразных эллиптических задач приведен в монографии [5]. Мы здесь кратко рассмотрим один из наиболее часто применяемых методов – итерационный метод Зейделя. Разрешим каждое из уравнений (6.40) относительно $u_{i,j}$ и запишем систему в виде, удобном для итераций [11].

$$u_{i,j} = \frac{1}{4}(u_{i+1,j} + u_{i-1,j} + u_{i,j+1} + u_{i,j-1} - d_{i,j}) \quad (6.42)$$

$$(i, j = 1, 2, \dots, n-1).$$

Зададим какое-либо начальное приближение к решению во всех внутренних точках, например, $u_{i,j}^{(0)} = 0$. Присвоим заданные значения $u_{i,j} = \varphi(x_i, y_j)$ для всех граничных точек. Построим алгоритм метода Зейделя для системы (6.42) в виде

$$u_{i,j}^{(k+1)} = \frac{1}{4}(u_{i+1,j}^{(k+1)} + u_{i-1,j}^{(k+1)} + u_{i,j+1}^{(k)} + u_{i,j-1}^{(k)} - d_{i,j}). \quad (6.43)$$

Верхний индекс в (6.43) указывает номер итерации. Напомним, что в методе Зейделя полученные на данной итерации новые значения переменных сразу используются для вычислений на этой итерации. Легко показать, что метод Зейделя в данном случае сходится, так как выполнены условия диагонального преобладания для матрицы исходной системы (6.40) (см. [11]). Действительно, для всех внутренних точек, не являющихся приграничными, элемент на главной диагонали равен 4 (коэффициент при $u_{i,j}$), а сумма остальных элементов в этой строке также равна 4. Достаточные условия сходимости метода Зейделя требуют, чтобы хотя бы в одном уравнении неравенство $|a_{ii}| \geq \sum_{i \neq j} |a_{ij}|$ было строгим. Очевидно, что в приграничных точках неравенство будет строгим. Например, в точке A на рис. 6.5 сумма элементов в системе (6.42), не стоящих на главной диагонали, равна 2, так как значения $u_{i,j}$ в точках E и D не входят в систему уравнений.

7. РЕШЕНИЕ ИНТЕГРАЛЬНЫХ УРАВНЕНИЙ

Интегральным уравнением называется уравнение, содержащее неизвестную функцию $y(x)$ под знаком интеграла. К интегральным уравнениям можно свести многие задачи математической физики с краевыми условиями. Кроме того, многие научные и инженерные задачи сводятся непосредственно к интегральным и к интегро-дифференциальным уравнениям. Под интегро-дифференциальными уравнениями понимают уравнения, в которых неизвестная функция $y(x)$ входит и под знаком интеграла, и в виде производных. К ним относятся, например, уравнения теплового излучения в задачах теплообмена, некоторые задачи оптики, радиотехники, гидроаэродинамики, квантовой механики и т.д. В этом разделе мы весьма кратко рассмотрим основные типы интегральных уравнений, постановки соответствующих задач и основные идеи их численного решения. С обзором интегральных уравнений и методов их решения можно ознакомиться в работе [1].

7.1. Основные понятия

Везде далее мы рассматриваем только **линейные интегральные уравнения**, в которых неизвестная функция $y(x)$ входит лишь в первой степени (линейно). Кроме того, рассматривается лишь одномерный случай, когда неизвестная функция y зависит лишь от одной независимой переменной x . Приведем некоторые виды интегральных

уравнений, которые чаще всего встречаются в практических приложениях и не требуют сложного математического аппарата для их объяснения.

Неоднородным интегральным уравнением Фредгольма второго рода называется уравнение

$$y(x) - \lambda \int_a^b K(x,s)y(s)ds = f(x), \quad x \in [a,b], \quad (7.1)$$

где $y(x)$ – неизвестная функция, подлежащая определению, $f(x)$ – правая часть уравнения, s – переменная интегрирования, $K(x,s)$ – ядро интегрального уравнения, λ – численный параметр. Если правая часть отсутствует [$f(x) = 0$], то уравнение (7.1) называется однородным уравнением

$$y(x) = \lambda \int_a^b K(x,s)y(s)ds, \quad x \in [a,b]. \quad (7.2)$$

Уравнением Фредгольма первого рода называется уравнение

$$\int_a^b K(x,s)y(s)ds = f(x), \quad x \in [a,b]. \quad (7.3)$$

Уравнение второго рода отличается тем, что в нем искомая функция $y(x)$ входит как под знаком интеграла, так и вне его, в то время как в уравнении первого рода $y(x)$ входит только под знаком интеграла. В уравнениях (7.1) – (7.3) обычно считают, что ядро $K(x,s)$ определено и ограничено на квадрате $a \leq x \leq b, a \leq s \leq b$.

Рассмотрим однородное уравнение Фредгольма второго рода (7.2). Это уравнение всегда имеет нулевое решение $y(x) = 0$, что легко проверить непосредственной подстановкой. Ненулевые решения $y(x) \neq 0$ однородного уравнения (7.2) существуют только для некоторых значений параметра λ . Те значения параметра λ , при которых однородное уравнение (7.2) имеет ненулевые решения, называют *собственными значениями* ядра $K(x,s)$ или данного интегрального уравнения, а соответствующие ненулевые решения однородного уравнения называют *собственными функциями*. Другими словами, если λ_i – собственное значение (7.2), а $\varphi_i(x)$ – собственная функция, то

$$\varphi_i(x) = \lambda_i \int_a^b K(x,s) \varphi_i(s) ds. \quad (7.4)$$

При решении интегральных уравнений (7.1) и (7.2) выполняются две задачи:

1) непосредственное нахождение точного или приближенного решения неоднородного уравнения (7.1) при заданном значении параметра λ и определенной функции $f(x)$;

2) отыскание собственных значений и собственных функций однородного уравнения (7.2). Связь между возможностью решения этих задач устанавливает следующая теорема.

Теорема Фредгольма. Если λ не является собственным значением однородного уравнения (7.2), то неоднородное уравнение (7.1) с непрерывной частью $f(x)$ имеет непрерывное единственное решение $y(x)$ при $a \leq x \leq b$; если же λ является собственным значением, то неоднородное уравнение (7.1) или вообще не имеет решений, или имеет их бесчисленное множество.

Отметим некоторые специальные виды ядра $K(x,s)$ интегрального уравнения, которые играют важную роль в практических приложениях. Вещественное ядро $K(x,s)$ называется симметричным, если выполняется условие $K(x,s) = K(s,x)$. Оно обладает следующими свойствами:

- а) симметричное ядро имеет хотя бы одно собственное значение;
- б) все собственные значения симметричного ядра – действительные числа;
- в) собственные функции симметричного ядра ортогональны.

Последнее означает, что если $\varphi_i(x)$ и $\varphi_j(x)$ являются собственными функциями одного и того же уравнения, то выполняется равенство

$$\int_a^b \varphi_i(x) \varphi_j(x) dx = 0 \quad \text{при } i \neq j. \quad (7.5)$$

Другим важным видом ядра являются вырожденные ядра. Ядро уравнения Фредгольма является вырожденным, если его можно представить в виде суммы членов, каждый из которых является произведением функций одной переменной

$$K(x,s) = \sum_{i=0}^n \varphi_i(x) \psi_i(s). \quad (7.6)$$

Отметим некоторые преимущества интегральных уравнений по сравнению с эквивалентными им дифференциальными задачами. Во-первых, интегральное уравнение содержит в себе полную постановку задачи, так как дополнительные условия (начальные или граничные) для него задавать не нужно. Они уже содержатся в нем. Например, задачу Коши для дифференциального уравнения

$$y' = f(x, y), \quad y(x_0) = y_0$$

легко представить в виде интегрального уравнения

$$y = y_0 + \int_{x_0}^x f(s, y(s)) ds,$$

в котором уже учтено начальное условие. Во-вторых, интегральные уравнения допускают простое обобщение на случай нескольких независимых переменных, что далеко не всегда просто сделать для дифференциальных уравнений.

7.2. Обзор численных методов

Одним из простых методов решения интегральных уравнений является **метод последовательных приближений**. Этот метод аналогичен методу простой итерации для одного нелинейного уравнения и заключается в построении последовательности функций $y_0(x), y_1(x), y_2(x), \dots, y_n(x)$, сходящейся к точному решению интегрального уравнения. Рассмотрим уравнение Фредгольма второго рода

$$y(x) = \lambda \int_a^b K(x, s) y(s) ds + f(x). \quad (7.7)$$

Зададим начальное приближение $y_0(x)$ искомой функции $y(x)$. Подставив $y_0(x)$ в правую часть уравнения, определим первое приближение по формуле

$$y_1(x) = \lambda \int_a^b K(x, s) y_0(s) ds + f(x).$$

Найденное значение $y_1(x)$ снова подставим в правую часть уравнения и определим $y_2(x)$

$$y_2(x) = \lambda \int_a^b K(x,s)y_1(s)ds + f(x).$$

В общем случае вычисление $(k + 1)$ -го приближения проводится по аналогичной формуле

$$y_{k+1}(x) = \lambda \int_a^b K(x,s)y_k(s)ds + f(x). \quad (7.8)$$

Можно показать, что при достаточно малом значении $|\lambda|$ и ограниченном ядре $K(x,s)$ этот итерационный процесс сходится. Достаточное условие сходимости имеет вид

$$M|\lambda|(b-a) < 1, \quad M = \max_{x,s} |K(x,s)|. \quad (7.9)$$

Другим методом является **метод аппроксимации ядра интегрального уравнения вырожденным ядром** (7.6). Подставляя (7.6) в интегральное уравнение (7.7), получим

$$y(x) = f(x) + \lambda \int_a^b \left| \sum_{i=0}^n \varphi_i(x) \psi_i(s) \right| y(s) ds = f(x) + \lambda \sum_{i=0}^n \varphi_i(x) C_i, \quad (7.10)$$

где C_i – постоянные коэффициенты, определяемые соотношением

$$C_i = \int_a^b \psi_i(s) y(s) ds, \quad (i = 1, 2, \dots, n). \quad (7.11)$$

Если теперь в (7.11) вместо $y(s)$ подставить его выражение (7.10), то для определения коэффициентов C_i получим систему линейных уравнений

$$C_i = \int_a^b \psi_i(s) f(s) ds + \lambda \int_a^b \psi_i(s) \sum_{j=1}^n C_j \varphi_j(s) ds, \quad (7.12)$$

которую можно проще записать в виде

$$C_i - \lambda \sum_{j=1}^n C_j \gamma_{ij} = f_i, \quad (i=1, 2, \dots, n), \quad (7.13)$$

где введены обозначения

$$f_i = \int_a^b \psi_i(s) f(s) ds, \quad \gamma_{ij} = \int_a^b \varphi_i(s) \psi_j(s) ds. \quad (7.14)$$

Если определитель системы (7.13) не равен нулю, то интегральное уравнение (7.7) имеет единственное решение, получаемое после определения величин C_i из уравнения (7.13).

Пример 17. Решить уравнение

$$y(x) = x^2 + \lambda \int_{-1}^1 (x+s) y(s) ds. \quad (7.15)$$

Решение. Ядро $K(x,s) = x + s$ является вырожденным. Из (7.15) получаем, если положить $\varphi_1(x) = x; \psi_1(s) = 1 = const; \varphi_2(x) = 1 = const; \psi_2(s) = s$;

$$y(x) = x^2 + \lambda (C_1 x + C_2), \quad (7.16)$$

где

$$C_1 = \int_{-1}^1 y(s) ds, \quad C_2 = \int_{-1}^1 s y(s) ds. \quad (7.17)$$

Подставляя (7.16) в интегралы (7.17), получим систему уравнений

$$C_1 = \frac{2}{3} + 2C_2\lambda, \quad C_2 = \frac{2}{3}\lambda C_1,$$

откуда сразу находим

$$C_1 = \frac{2/3}{1 - \frac{4}{3}\lambda^2}, \quad C_2 = \frac{(4\lambda)/9}{1 - \frac{4}{3}\lambda^2}. \quad (7.18)$$

Теперь по формуле (7.16) получаем решение интегрального уравнения [при условии, что знаменатель в (7.18) не равен нулю, или иначе $\lambda^2 \neq 3/4$]

$$y(x) = x^2 + \frac{2\lambda x + (4/3)\lambda^2}{3 - 4\lambda^2}. \quad (7.19)$$

При $\lambda^2 = \frac{3}{4}$ уравнение (7.15) решений не имеет.

Широкое применение находят также **квадратурные методы**, основанные на использовании различных формул численного интегрирования для вычислений определенных интегралов, входящих в интегральное уравнение. Обычно используются квадратурные методы трапеций, Симпсона или Гаусса. Пусть определенный интеграл вычисляется с помощью некоторой квадратурной формулы

$$\int_a^b F(x)dx = \sum_{i=1}^n A_i F(x_i), \quad (7.20)$$

где x_i ($i = 1, 2, \dots, n$) – абсциссы точек отрезка $[a, b]$, делящие этот отрезок на равные элементарные отрезки $[x_{i-1}, x_i]$. Коэффициенты A_i легко определяются для каждой квадратурной формулы. Например, для метода трапеций

$$A_1 = A_n = h/2, \quad A_2 = A_3 = \dots = A_{n-1} = h;$$

для составной формулы Симпсона

$$A_1 = A_{2m+1} = h/3; A_2 = A_4 = \dots = A_{2m} = 4h/3; A_3 = A_5 = \dots = A_{2m-1} = 2h/3.$$

Пусть задано интегральное уравнение Фредгольма второго рода

$$y(x) - \lambda \int_a^b K(x,s)y(s)ds = f(x), \quad x \in [a,b]. \quad (7.21)$$

Беря узлы $x_i \in [a,b]$ и $s_j \in [a,b]$ и вводя обозначения:

$$y(x_i) = y_i, \quad y(s_j) = y_j, \quad K(x_i, s_j) = K_{ij}, \quad f(x_i) = f_i, \quad (i, j = 1, 2, \dots, n),$$

на основании формулы (7.20) получим

$$y_i - \lambda \sum_{j=1}^n A_j K_{ij} y_j = f_i, \quad (i = 1, 2, \dots, n). \quad (7.22)$$

Уравнения (7.22) представляют собой систему линейных уравнений относительно сеточной функции y_i , которая имеет решение, если λ не равны собственным значениям однородного уравнения [см. формулу (7.2)].

ЗАДАНИЯ ДЛЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ

1. Составьте логическую схему базы знаний по теме юниты.

2. Решите самостоятельно следующие задачи, выбрав свой вариант. Во всех задачах величина S – номер студента в журнале и g – номер группы. Задав величины S и g , студент определяет условия задачи своего варианта.

2.1. Задано дифференциальное уравнение

$$y'' + xy' + y = 0,3Sx^2 + 0,2(S + g)$$

и начальные условия $y(0) = 0,2g$; $y'(0) = 0$.

а) Свести задачу Коши к системе уравнений первого порядка.

б) Решить данную задачу Коши на интервале $[0; 0,4]$ с двумя различными шагами $h = 0,1$ и $h = 0,2$ методом Эйлера. Полученные результаты $y(0)$, $y(0,2)$ и $y(0,4)$ для шага $h = 0,2$ и $y(0)$, $y(0,1)$, $y(0,2)$, $y(0,3)$ и $y(0,4)$ для шага $h = 0,1$ представить в виде таблицы и изобразить на графике.

в) В точках $y(0,2)$ и $y(0,4)$ провести оценку погрешности по методу Рунге, получить уточненные значения решения и представить их на графике.

г) Для сеточной функции $y(x)$, вычисленной с шагом $h = 0,1$, вычислить интеграл

$$\int_0^{0,4} y(x) dx,$$

используя составную квадратурную формулу Симпсона.

д) Для сеточной функции, полученной в п. б) с шагом $h = 0,2$, получить численно значения производных $y'(0)$, $y'(0,2)$, $y'(0,4)$, используя формулы второго порядка точности.

2.2. Задано дифференциальное уравнение

$$y'' + xy' + y = 0,3Sx^2 + 0,2(S + g)$$

и граничные условия $y(0) = 0,2g$; $y'(0,4) = 0,08S$.

а) Решить краевую задачу конечно-разностным методом, полагая $n = 2$ ($h = 0,2$).

б) По формулам линейной интерполяции получить значения $y(0,1)$ и $y(0,3)$ с помощью сеточной функции $y(0)$, $y(0,2)$, $y(0,4)$.

ТРЕНИНГ УМЕНИЙ

1. Пример выполнения упражнения тренинга на умение № 1

Задание

Функция $y = f(x)$ задана таблицей

x	0	0,2	0,4
y	0	0,2214	0,4918

С помощью линейной интерполяции найти $f(0,3)$.

С помощью квадратичной интерполяции найти $f(0,15)$.

Решение

№ п/п	Алгоритм	Конкретное соответствие данной ситуации предложенному алгоритму
1.	По заданному значению x_* определить узловые точки x_i, x_{i+1} , для которых $x_* \in [x_i, x_{i+1}]$.	Запишем формулу линейной интерполяции (1.22) $y = y_0 + \frac{y_1 - y_0}{x_1 - x_0}(x - x_0)$ Необходимо найти функцию при значении аргумента $x = 0,3$, который находится на интервале $[0,2; 0,4]$. Поэтому примем $x_0 = 0,2$; $x_1 = 0,4$; $y_0 = 0,2214$; $y_1 = 0,4918$. $y(0,3) = 0,2214 + \frac{0,4918 - 0,2214}{0,4 - 0,2}(0,3 - 0,2) =$ $= 0,2214 + 0,1362 = 0,3566.$
2.	Построить линейный интерполяционный многочлен [см. формулу (1.22)] $L_1(x)$ и вычислить $L_1(x_*)$.	В данном случае надо найти функцию $f(x)$ при $x = 0,15$. Примем $x_0 = 0$; $x_1 = 0,2$; $y_0 = 0$; $y_1 = 0,2214$. По формуле линейной интерполяции $y(0,15) = 0 + \frac{0,2214 - 0}{0,2 - 0}(0,15 - 0) = 0,1661.$

№ п/п	Алгоритм	Конкретное соответствие данной ситуации предложенному алгоритму
3.	Построить квадратный интерполяционный многочлен $L_2(x)$ [см. формулу (1.23)] и вычислить $L_2(x_*)$.	Запишем формулу квадратичной интерполяции $y + y_0 + \frac{y_1 - y_0}{h}(x - x_0) + \frac{y_2 - 2y_1 + y_0}{2h^2}(x - x_0)(x - x_1).$ Здесь $h = x_1 - x_0 = 0,2$; $x_2 = 0,4$; $y_2 = 0,4918$. Точки x_0, x_1, y_0, y_1 берутся те же, что рассматривались выше, при линейной интерполяции $y = 0 + \frac{0,2214 - 0}{0,2 - 0} \cdot 0,15 + \frac{0,4918 - 2 \cdot 0,2214 + 0}{2 \cdot (0,2)^2} \cdot (0,15 - 0)(0,15 - 0,2) =$ $= 0,1661 - 0,00459 = 0,1615.$ Первые два члена в правой части формулы квадратичной интерполяции дают результат линейной интерполяции, а последний член – поправка, даваемая квадратичной интерполяцией.

Решите самостоятельно:

Задание 1.1

Функция $y = f(x)$ задана таблицей

x	0,5	0,6	0,7	0,8
y	1,6487	1,8221	2,0138	2,2255

С помощью линейной интерполяции найти $f(0,73)$.

С помощью квадратичной интерполяции найти $f(0,52)$.

Задание 1.2

Функция $y = f(x)$ задана таблицей

x	1	1,1	1,2
y	0,0000	0,0953	0,1823

С помощью линейной интерполяции найти $f(1,13)$.

С помощью квадратичной интерполяции найти $f(1,09)$.

Задание 1.3

Функция $y = f(x)$ задана таблицей

x	0,6	0,8	1,0
y	0,6841	1,0296	1,5574

С помощью линейной интерполяции найти $f(0,95)$.

С помощью квадратичной интерполяции найти $f(0,68)$.

2. Пример выполнения упражнения тренинга на умение № 2

Задание

Функция $y = f(x)$ задана таблицей

i	0	1	2	3	4
x_i	0	0,25	0,5	0,75	1,0
y_i	1,0000	1,2840	1,6487	2,1170	2,7183

По формуле Симпсона вычислить значение интеграла

$$S = \int_0^1 f(x)$$

а) с шагом $h = 0,5$, используя три точки разбиения отрезка $[0, 1]$: при $x = 0; 0,5; 1$;

б) с шагом $h = 0,25$, используя все пять точек разбиения отрезка $[0, 1]$: при $x = 0; 0,25; 0,5; 0,75; 1$.

Вычислить уточненное значение интеграла $\bar{S}$ с помощью метода Рунге, используя полученные значения интеграла с шагами $h = 0,5$ и $h = 0,25$.

Решение

№ п/п	Алгоритм	Конкретное соответствие данной ситуации предложенному алгоритму
1.	Разделить отрезок интегрирования на n и $2n$ равных частей (с шагами h и $h/2$).	При вычислении интеграла рассмотрим два различных разбиения отрезка интегрирования $[0, 1]$. При первом разбиении возьмем два элементарных отрезка ($n = 2$): $[0; 0,5]$, $[0,5; 1]$. В этом случае $h = 0,5$. Точками разбиения будут точки $x_0 = 0$; $x_1 = 0,5$; $x_2 = 1$. Соответствующие значения функции равны: $y_0 = 1$; $y_1 = 1,6487$; $y_2 = 2,7183$. Во втором разбиении возьмем четыре элементарных отрезка ($n = 4$): $[0; 0,25]$, $[0,25; 0,5]$, $[0,5; 0,75]$, $[0,75; 1]$. Теперь $h = 0,25$, т.е. шаг в два раза меньше, чем в первом случае. Точками разбиения теперь будут точки: $x_0 = 0$; $x_1 = 0,25$; $x_2 = 0,5$; $x_3 = 0,75$; $x_4 = 1$. Соответствующие значения функции y_0, y_1, y_2, y_3, y_4 указаны в таблице.
2.	По формуле Симпсона вычислить значения интеграла S_h с шагом h и $S_{h/2}$ с шагом $h/2$ [см. формулу (2.25)].	Вычислим значения интеграла $S_{0,5}$ с шагом $h = 0,5$ и $S_{0,25}$ с шагом $h = 0,25$ по формуле Симпсона (2.25). $S_{0,5} = \frac{h}{3}(y_0 + 4y_1 + y_2) =$ $= \frac{0,5}{3}(1 + 1,6487 \cdot 4 + 2,7183) = 1,7189.$

№ п/п	Алгоритм	Конкретное соответствие данной ситуации предложенному алгоритму
		$S_{0,25} = \frac{h}{3} (y_0 + 4y_1 + 2y_2 + 4y_3 + y_4) = \frac{0,25}{3} \cdot$ $\cdot (1 + 4 \cdot 1,2840 + 2 \cdot 1,6487 + 4 \cdot 2,1170 + 2,7183) =$ $= 1,7183.$
3.	Вычислить уточненное значение интеграла с помощью метода Рунге [формула (2.45)].	$\bar{S} = S_{0,25} + \frac{S_{0,25} - S_{0,5}}{15} =$ $= 1,7183 + \frac{1,7183 - 1,7189}{15} = 1,71826.$

Решите самостоятельно:

Задание 2.1

Вычислить интеграл $\int_0^1 e^{-x^2} dx$

а) вычислить значения подынтегральной функции $f(x) = e^{-x^2}$ в узловых точках $x_0 = 0$; $x_1 = 0,25$; $x_2 = 0,5$; $x_3 = 0,75$; $x_4 = 1$;

б) по формуле Симпсона вычислить значение интеграла с шагом $h = 0,5$, взяв три узловые точки $x = 0$; $x = 0,5$; $x = 1$;

в) по формуле Симпсона вычислить значения интеграла с шагом $h = 0,25$, используя полученные в п. а значения функции $f(x_0)$, $f(x_1)$, $f(x_2)$, $f(x_3)$, $f(x_4)$;

г) вычислить уточненное значение интеграла по методу Рунге, используя значения интеграла, полученные с шагами $h = 0,5$ и $h = 0,25$.

Задание 2.2 ✓

Функция $y = f(x)$ задана таблицей:

x	0,2	0,4	0,6
y	0,1987	0,3694	0,5646

По формуле трапеций вычислить интеграл $\int_0^{0,6} f(x) dx$

✓ Задание 2.3

Вычислить методом прямоугольников интеграл $\int_0^1 e^{x^2} dx$, взяв $n = 4$
($h = 0,25$).

3. Пример выполнения упражнения тренинга на умение № 3

Задание

Функция $y = f(x)$ задана таблицей

i	0	1	2	3
x_i	0	0,2	0,4	0,6
y_i	1,0000	0,9801	0,9211	0,8253

Вычислить первую и вторую производные во внутренних точках $x_1 = 0,2$ и $x_2 = 0,4$ с помощью центральных разностей.

Вычислить первую и вторую производные $f'(x)$ и $f''(x)$ в граничных точках $x_0 = 0$, $x_3 = 0,6$.

Решение

№ п/п	Алгоритм	Конкретное соответствие данной ситуации предложенному алгоритму
1.	Вычислить первую и вторую производные во внутренних точках с помощью центральных разностей [см. формулы (3.18) и (3.20)].	Вычислим первые производные в точках x_1 и x_2 по формуле (3.18): $f'(x_1) = \frac{y_2 - y_0}{2h} = \frac{0,9211 - 1}{2 \cdot 0,2} = -0,1973;$ $f'(x_2) = \frac{y_3 - y_1}{2h} = \frac{0,8253 - 0,9801}{2 \cdot 0,2} = -0,3870.$ Вторые производные в точках x_1 и x_2 вычислим по формуле (3.20): $f''(x_1) = \frac{y_2 - 2y_1 + y_0}{h^2} =$ $= \frac{0,9211 - 2 \cdot 0,9801 + 1}{(0,2)^2} = -0,9775;$ $f''(x_2) = \frac{y_3 - 2y_2 + y_1}{h^2} =$ $= \frac{0,8253 - 2 \cdot 0,9211 + 0,9801}{(0,2)^2} = -0,92.$

№ п/п	Алгоритм	Конкретное соответствие данной ситуации предложенному алгоритму
2.	Вычислить первую и вторую производные в граничных точках с погрешностями $O(h)$ и $O(h^2)$ [см. формулы (3.16), (3.27), (3.29), (3.31), (3.33),].	Вычислим первую и вторые производные в левой граничной точке $x_0 = 0$ по формулам (3.15), (3.27) и (3.31). По формуле (3.15), имеющей погрешность $O(h)$, $f'(0) = \frac{y_1 - y_0}{h} = \frac{0,9801 - 1}{0,2} = -0,0995.$ По формуле (3.27), имеющей погрешность $O(h^2)$, $f'(0) = \frac{-3y_0 + 4y_1 - y_2}{2h} =$ $= \frac{-3 \cdot 1 + 4 \cdot 0,9801 - 0,9211}{2 \cdot 0,2} = -0,00175.$ Вычислим теперь $f''(0)$ по формуле (3.31) с погрешностью $O(h)$ $f''(0) = \frac{y_0 - 2y_1 + y_2}{h^2} =$ $= \frac{1 - 2 \cdot 0,9801 + 0,9211}{(0,2)^2} = -0,9775.$ Аналогично вычислим первую и вторую производные в правой граничной точке $x_3 = 0,6$ по формулам (3.16), (3.29) и (3.33): $f'(0,6) = \frac{y_3 - y_2}{h} = \frac{0,8253 - 0,9211}{0,2} = -0,4790;$ $f'(0,6) = \frac{3y_3 - 4y_2 + y_1}{2h} =$ $= \frac{3 \cdot 0,8253 - 4 \cdot 0,9211 - 0,9801}{2 \cdot 0,2} = -0,5710;$ $f''(0,6) = \frac{y_3 - 2y_2 + y_1}{h^2} =$ $= \frac{0,8253 - 2 \cdot 0,9211 + 0,9801}{(0,2)^2} = -0,9200.$

Решите самостоятельно:

✓ Задание 3.1

Функция $y = f(x)$ задана таблицей:

x	0,5	0,6	0,7
y	0,4794	0,5646	0,6442

Вычислить первую и вторую производные при всех значениях аргумента.

Учтя, что функция, заданная в таблице, есть $y = \sin x$, оценить погрешность полученных производных, сравнивая с точным решением $y' = \cos x$, $y'' = -\sin x$.

разности-х эквивалент
разв. членов-от.

Задание 3.2

Функция $y = f(x)$ задана таблицей:

x	0	0,4	0,8
y	0	0,4228	1,0296

Вычислить значения $f'(x)$ и $f''(x)$ в граничных точках $x=0$,
 $x_2=0,8$. $y_2=1,0296$ $x_1=0,4$ $y_1=0,4228$ $x_0=0$ $y_0=0$ $h=0,4$

$$f'(0,8) = \frac{y_2 - y_1}{h} = \frac{1,0296 - 0,4228}{0,4} = 1,517 \quad O(h)$$

$$f'_h(0,8) = \frac{3y_2 - 4y_1 + y_0}{2h} = \frac{3 \cdot 1,0296 - 4 \cdot 0,4228 + 0}{2 \cdot 0,4} = \frac{1,3876}{0,8}$$

$$= 1,7347$$

$$f''(0,8) = \frac{y_0 - 2y_1 + y_2}{h^2} = \frac{0 - 0,8456 + 1,0296}{0,16} = 1,131$$

Задание 3.3

Функция $y = f(x)$ задана формулой $y = \sqrt{\frac{e^x + \sin x}{x^2 + 1}}$. С помощью

калькулятора вычислить значения этой функции в точках $x = 0$; $x = 0,1$; $x = 0,2$. Вычислить при этих значениях x значения первой производной $f'(0)$, $f'(0,1)$, $f'(0,2)$.

4. Пример выполнения упражнения тренинга на умение № 4

Задание

Для дифференциального уравнения

$$y'' + xy' + y = 3x^2 + 2$$

найти решение задачи Коши $y(0) = 0$, $y'(0) = 0$, сделав с шагом $h = 0,2$ два шага методом Эйлера.

Решение

№ п/п	Алгоритм	Конкретное соответствие данной ситуации предложенному алгоритму
1.	Записать уравнение в виде системы двух уравнений первого порядка.	Запишем уравнение в виде системы двух уравнений первого порядка. Обозначим $z = y'$. Получим систему: $\begin{cases} y' = z \\ z' = 3x^2 + 2 - xz - y \end{cases}$
2.	Записать начальные условия для системы.	Начальные условия примут вид $z(0) = 0$, $y(0) = 0$. Взяв $h = 0,2$, сделаем первый шаг методом Эйлера. Обозначим $y(0) = y_0$, $y(0,2) = y_1$, $y(0,4) = y_2$, $z(0) = z_0$, $z(0,2) = z_1$, $z(0,4) = z_2$.
3.	Найти решение системы, сделав два шага методом Эйлера или методом Эйлера с пересчетом [см. формулы (5.26) и (5.29), (5.30)].	Запишем формулы (5.47), определяющие один шаг метода Эйлера в данном конкретном случае. Для общей системы двух уравнений первого порядка с начальными условиями $\begin{cases} y' = f_1(x, y, z) & y(x_0) = y_0 \\ z' = f_2(x, y, z) & z(x_0) = z_0 \end{cases}$ один шаг метода Эйлера можно записать в виде $\begin{cases} y_1 = y_0 + hf_1(x_0, y_0, z_0) \\ z_1 = z_0 + hf_2(x_0, y_0, z_0) \\ x_1 = x_0 + h \end{cases}$

№ п/п	Алгоритм	Конкретное соответствие данной ситуации предложенному алгоритму																																
		В нашем случае $x_0 = 0, y_0 = 0, z_0 = 0, \quad f_1(x, y, z) = z,$ $f_2(x, y, z) = 3x^2 + 2 - xz - y.$ Следовательно, $\begin{cases} y_1 = 0 + 0,2 \cdot 0 = 0 \\ z_1 = 0 + 0,2(3 \cdot 0^2 + 2 - 0 \cdot 0 - 0) = 0,4 \\ x_1 = 0 + 0,2 = 0,2 \end{cases}$ Примем значения $y_1 = 0$; $z_1 = 0,4$ и $x_1 = 0,2$ за исходные значения и сделаем еще один шаг метода Эйлера $\begin{cases} y_2 = y_1 + h z_1 = 0 + 0,2 \cdot 0,4 = 0,08 \\ z_2 = z_1 + h(3x_1^2 + 2 - x_1 z_1 - y_1) = \\ = 0,4 + 0,2(3 \cdot (0,2)^2 + 2 - 0,2 \cdot 0,4 - 0) = \\ = 0,4 + 0,2(0,12 + 2 - 0,08) = 0,808 \\ x_2 = x_1 + h = 0,2 + 0,2 = 0,4 \end{cases}$ Точным решением задачи Коши для исходного уравнения второго порядка является функция $y = x^2$, что легко проверить, подставив $y = x^2$ в исходное уравнение и в начальные условия. В приводимой ниже таблице даются полученные методом Эйлера значения y_1, y_2 и z_1, z_2, их точные значения $\bar{y}_i$ и $\bar{z}_i$ и погрешности решения методом Эйлера $y_i - \bar{y}_i$, $z_i - \bar{z}_i$ <table><tr><th>i</th><th>x_i</th><th>y_i</th><th>z_i</th><th>$\bar{y}_i$</th><th>$\bar{z}_i$</th><th>$y_i - \bar{y}_i$</th><th>$z_i - \bar{z}_i$</th></tr><tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr><tr><td>1</td><td>0,2</td><td>0</td><td>0,4</td><td>0,04</td><td>0,4</td><td>0,04</td><td>0</td></tr><tr><td>2</td><td>0,4</td><td>0,08</td><td>0,808</td><td>0,160</td><td>0,800</td><td>0,08</td><td>0,008</td></tr></table>	i	x_i	y_i	z_i	$\bar{y}_i$	$\bar{z}_i$	$ y_i - \bar{y}_i $	$ z_i - \bar{z}_i $	0	0	0	0	0	0	0	0	1	0,2	0	0,4	0,04	0,4	0,04	0	2	0,4	0,08	0,808	0,160	0,800	0,08	0,008
i	x_i	y_i	z_i	$\bar{y}_i$	$\bar{z}_i$	$ y_i - \bar{y}_i $	$ z_i - \bar{z}_i $																											
0	0	0	0	0	0	0	0																											
1	0,2	0	0,4	0,04	0,4	0,04	0																											
2	0,4	0,08	0,808	0,160	0,800	0,08	0,008																											

Решите самостоятельно:

✓ **Задание 4.1**

Для дифференциального уравнения

$$y'' + xy = x^2$$

найти решение задачи Коши $y(1) = 0$, $y'(1) = 0,5$, сведя уравнение к системе двух уравнений первого порядка и сделав три шага методом Эйлера с шагом $h = 0,1$.

Задание 4.2

Для задачи Коши $y' = x^2 + y$, $y(0) = 1$ сделать два шага методом Эйлера с шагом $h = 0,3$.

$$x_0 = 0; y_0 = 1$$

$$y_1 = y_0 + h \cdot f(x_0; y_0)$$

$$f(x, y) = x^2 + y$$

$$f(x_0; y_0) = 0^2 + 1 = 1$$

$$y_1 = 1 + 0,3 \cdot 1 = 1,3$$

$$y_2 = y_1 + h \cdot f(x_1; y_1)$$

$$x_1 = x_0 + h = 0,3$$

$$f(x_1; y_1) = 0,3^2 + 1,3 = 1,39$$

$$y_2 = 1,3 + 0,3 \cdot 1,39 = 1,717$$

Задание 4.3

Для задачи Коши $y' = x^2 + y$, $y(0) = 1$ сделать два шага с $h = 0,1$ методом Эйлера с пересчетом.

5. Пример выполнения упражнения тренинга на умение № 5

Задание

Решить конечно-разностным методом краевую задачу

$$y'' - 2xy' - 2y = -4x; \quad y(0) - y'(0) = 0; \quad y(1) = 3,718; \quad x \in [0,1],$$

разбив отрезок $[0, 1]$ на четыре равных интервала ($n = 4$).

Решение

№ п/п	Алгоритм	Конкретное соответствие данной ситуации предложенному алгоритму
1.	Написать конечно-разностные уравнения для внутренних точек отрезка.	Введем сетку $x_0 = 0; x_1 = 0,25; x_2 = 0,5; x_3 = 0,75; x_4 = 1$. Обозначим искомые значения сеточной функции y_0, y_1, y_2, y_3, y_4. Запишем прежде всего аппроксимацию дифференциального уравнения разностным во внутренних точках отрезка x_1, x_2 и x_3. При $x_1 = 0,25$, используя для y'' и y' формулы (3.18) и (3.20), получим $\frac{y_2 - 2y_1 + y_0}{(0,25)^2} - 2 \cdot 0,25 \frac{y_2 - y_0}{2 \cdot 0,25} - 2y_1 = -4 \cdot 0,25.$ После простых преобразований это уравнение примет вид $17y_0 - 34y_1 + 15y_2 = -1. \quad (T1)$ Запишем теперь разностное уравнение при $x_2 = 0,5$ $\frac{y_3 - 2y_2 + y_1}{(0,25)^2} - 2 \cdot 0,5 \frac{y_3 - y_1}{2 \cdot 0,25} - 2y_2 = -4 \cdot 0,5$ или после упрощений $18y_1 - 34y_2 + 14y_3 = -2. \quad (T2)$ При $x_3 = 0,75$ разностное уравнение имеет вид $\frac{y_4 - 2y_3 + y_2}{(0,25)^2} - 2 \cdot 0,75 \frac{y_4 - y_2}{2 \cdot 0,25} - 2y_3 = -4 \cdot 0,75$

№ п/п	Алго- ритм	Конкретное соответствие данной ситуации предложенному алгоритму
		или после упрощений $19y_2 - 34y_3 + 13y_4 = -3. \quad (T3)$
2.	Напи- сать конеч- но- разно- стную аппро- ксима- цию для гра- ничных усло- вий.	Запишем теперь разностные уравнения, аппроксимирующие граничные условия. Граничное условие $y(0) - y'(0) = 0$ заменим разностным соотношением $y_0 - \frac{y_1 - y_0}{0,25} = 0$, где величина $y'(0)$ аппроксимируется с помощью правой разности по формуле (3.15). После приведения подобных членов уравнение примет вид $5y_0 - 4y_1 = 0. \quad (T4)$ На правом конце отрезка граничное условие имеет простой вид $y_4 = 3,718. \quad (T5)$ Запишем теперь полученную систему уравнений (T1)–(T5) $\begin{cases} 5y_0 - 4y_1 & = 0 \\ 17y_0 - 34y_1 + 15y_2 & = -1 \\ 18y_1 - 34y_2 + 14y_3 & = -2 \\ 19y_2 - 34y_3 + 13y_4 & = -3 \\ y_4 & = 3,718 \end{cases} \quad (T6)$
3.	Решить полу- чен- ную систе- му ме- тодом Гаус- са.	Систему (T6) можно решать как методом Гаусса, так и итерационным методом Зейделя, так как матрица коэффициентов обладает свойством диагонального преобладания (см. [11]). Кроме того, эту систему можно решать методом прогонки [11], который является в сущности удобной реализацией метода Гаусса для системы уравнений с трехдиагональной матрицей. (Кстати, матрица системы (T6) является трехдиагональной). Но систему (T6) можно свести к системе трех уравнений с тремя неизвестными. Для этого выразим y_0 через y_1 из первого уравнения и подставим это значение вместо y_0 во второе уравнение. В четвертом уравнении системы (T6) заменим y_4 на его значение $y_4 = 3,718$. В результате получим систему уравнений

№ п/п	Алгоритм	Конкретное соответствие данной ситуации предложенному алгоритму																								
		$\begin{cases} -20,4y_1 + 15y_2 & = -1 \\ 18y_1 - 34y_2 + 14y_3 & = -2 \\ 19y_2 - 34y_3 & = -51,334 \end{cases} \quad (T7)$ Систему (Т7) легко решить методом Гаусса. Не приводя выкладок, выпишем окончательное решение системы в виде таблицы с учетом значений в граничных точках <table><tr><th>I</th><th>0</th><th>1</th><th>2</th><th>3</th><th>4</th></tr><tr><td>x_i</td><td>0</td><td>0,25</td><td>0,5</td><td>0,75</td><td>1</td></tr><tr><td>y_i</td><td>1,1310</td><td>1,4138</td><td>1,8561</td><td>2,5471</td><td>3,7183</td></tr><tr><td>$\tilde{y}_i$</td><td>1</td><td>1,3145</td><td>1,7840</td><td>2,5051</td><td>3,7183</td></tr></table> Для сравнения в таблице в последней строке дается точное решение этой задачи $\tilde{y}_i$ в узлах выбранной сетки. Точное решение рассматриваемой краевой задачи имеет вид $\tilde{y}(x) = x + e^{x^2}.$ Если учесть, что была выбрана довольно грубая сетка, то полученные погрешности можно считать не слишком большими. Даже на этой грубой сетке можно получить еще более точные результаты, если более точно аппроксимировать левое граничное условие $y(0) - y'(0) = 0$. Разностное уравнение (Т4) имеет погрешность $O(h)$, так как для аппроксимации $y'(0)$ использовалась правая разность (3.15), имеющая погрешность $O(h)$. Более точные результаты можно получить, если для $y'(0)$ использовать аппроксимацию (3.27), имеющую погрешность $O(h^2)$ $y'(0) \approx \frac{-3y_0 + 4y_1 - y_2}{2h} = \frac{-3y_0 + 4y_1 - y_2}{2 \cdot 0,25}.$ Тогда левое граничное условие запишется в виде $y_0 = \frac{-3y_0 + 4y_1 - y_2}{0,5}$	I	0	1	2	3	4	x_i	0	0,25	0,5	0,75	1	y_i	1,1310	1,4138	1,8561	2,5471	3,7183	$\tilde{y}_i$	1	1,3145	1,7840	2,5051	3,7183
I	0	1	2	3	4																					
x_i	0	0,25	0,5	0,75	1																					
y_i	1,1310	1,4138	1,8561	2,5471	3,7183																					
$\tilde{y}_i$	1	1,3145	1,7840	2,5051	3,7183																					

№ п/п	Алгоритм	Конкретное соответствие данной ситуации предложенному алгоритму																																			
		или после упрощений $7y_0 - 8y_1 + 2y_2 = 0. \quad (T8)$ Записав в системе (Т6) вместо первого уравнения уравнение (Т8), мы получим систему уравнений, которая дает более точное решение исходной краевой задачи $\begin{cases} 7y_0 - 8y_1 + 2y_2 = 0 \\ 17y_0 - 34y_1 + 15y_2 = -1 \\ 18y_1 - 34y_2 + 14y_3 = -2 \\ 19y_2 - 34y_3 = -51,334 \end{cases} \quad (T9)$ Студентам предлагается самостоятельно решить эту систему методом Гаусса и сравнить результат с двумя последними строками в таблице. Из первых двух уравнений можно исключить y_0, и тогда система примет вид $\begin{cases} -14,5714y_1 + 10,1429y_2 = -1 \\ 18y_1 - 34y_2 + 14y_3 = -2 \\ 19y_2 - 34y_3 = -51,334 \end{cases} \quad (T10)$ Решение этой системы методом Гаусса дает следующие значения: $y_1 = 1,3118; y_2 = 1,7860; y_3 = 2,5079.$ Сведем все полученные решения в таблицу <table><tr><th>i</th><th>0</th><th>1</th><th>2</th><th>3</th><th>4</th><th></th></tr><tr><th>x_i</th><td>0</td><td>0,25</td><td>0,5</td><td>0,75</td><td>1</td><td></td></tr><tr><th>$y_{i,1}$</th><td>1,1310</td><td>1,4138</td><td>1,8561</td><td>2,5471</td><td>3,7183</td><td>погрешность $O(h)$ для левой границы</td></tr><tr><th>$y_{i,2}$</th><td>0,9889</td><td>1,3118</td><td>1,7860</td><td>2,5079</td><td>3,7183</td><td>погрешность $O(h^2)$ для левой границы</td></tr><tr><th>$\tilde{y}_i$</th><td>1</td><td>1,3145</td><td>1,7840</td><td>2,5051</td><td>3,7183</td><td>точное решение</td></tr></table> Видно, что более точная аппроксимация $O(h^2)$ левого граничного условия дает гораздо более точное решение.	i	0	1	2	3	4		x_i	0	0,25	0,5	0,75	1		$y_{i,1}$	1,1310	1,4138	1,8561	2,5471	3,7183	погрешность $O(h)$ для левой границы	$y_{i,2}$	0,9889	1,3118	1,7860	2,5079	3,7183	погрешность $O(h^2)$ для левой границы	$\tilde{y}_i$	1	1,3145	1,7840	2,5051	3,7183	точное решение
i	0	1	2	3	4																																
x_i	0	0,25	0,5	0,75	1																																
$y_{i,1}$	1,1310	1,4138	1,8561	2,5471	3,7183	погрешность $O(h)$ для левой границы																															
$y_{i,2}$	0,9889	1,3118	1,7860	2,5079	3,7183	погрешность $O(h^2)$ для левой границы																															
$\tilde{y}_i$	1	1,3145	1,7840	2,5051	3,7183	точное решение																															

Решите самостоятельно:

Задание 5.1

Решить конечно-разностным методом краевую задачу

$$y'' - x^2 y = 1; \quad y(-1) = 0; \quad y(1) = 0; \quad x \in [-1, 1];$$

разбив отрезок $[-1, 1]$ на четыре равных интервала ($n = 4, h = 0,5$).

Введем сетку $x_0 = -1; x_1 = -0,5; x_2 = 0; x_3 = 0,5; x_4 = 1$

$$x_1 = -0,5$$

$$y_0 = 0; y_4 = 0$$

$$\frac{y_2 - 2y_1 + y_0}{h^2} - x_1^2 y_1 = 1$$

$$\frac{y_2 - 2y_1 + 0}{0,25} - 0,25 y_1 = 1$$

$$y_2 - 2y_1 - 0,0625 y_1 = 0,25 \cdot 16$$

$$16y_2 - 32y_1 - y_1 = 4$$

$$16y_2 - 33y_1 = 4$$

$$x_1 = -0,5; 16y_2 - 33y_1 = 4$$

$$x_2 = 0$$

$$\frac{y_3 - 2y_2 + y_1}{h^2} - x_2^2 y_2 = 1$$

$$\frac{y_3 - 2y_2 + y_1}{0,25} = 1$$

$$y_3 + 2y_2 + y_1 = 0,25 \cdot 4$$

$$x_2 = 0 \quad 4y_3 - 8y_2 + 4y_1 = 1$$

$$x_3 = 0,5$$

$$\frac{y_4 - 2y_3 + y_2}{h^2} - x_3^2 y_3 = 1 \cdot 0,25$$

$$y_2 - 2y_3 - 0,25 \cdot 0,25 \cdot y_3 = 0,25$$

$$16y_2 - 32y_3 - y_3 = 4$$

$$x_3 = 0,5 \quad 16y_2 - 33y_3 = 4$$

$$16y_2 - 33y_3 = 4$$

$$\begin{cases} 4y_3 - 8y_2 + 4y_1 = 1 \\ 16y_2 - 33y_3 = 4 \\ -33y_3 + 16y_2 = 4 \end{cases}$$

Задание 5.2

Решить конечно-разностным методом краевую задачу

$$y'' - y = x; \quad y'(0) = y'(1) = 0; \quad x \in [0,1],$$

разбив отрезок $[0, 1]$ на три равных интервала ($n = 3, h = 1/3$).

Задание 5.3

Решить конечно-разностным методом краевую задачу

$$y'' = x; \quad y'(0) = 0, y'(1) = 0,5; \quad x \in [0,1],$$

разбив отрезок $[0, 1]$ на три равных интервала ($n = 3, h = 3$). Непосредственным интегрированием дифференциального уравнения получите точное решение. Сравните точное решение с результатами решения задачи конечно-разностным методом.